Normal Form for Fermi-Pasta-Ulam Chains

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Andreas Henrici

von

Basel BS

Promotionskomitee

Prof. Dr. Thomas Kappeler (Vorsitz)
Prof. Dr. Jürgen Pöschel

Zürich, 2008

Zusammenfassung

In dieser Arbeit untersuchen wir die Normalformtheorie eindimensionaler Fermi-Pasta-Ulam-Ketten. Wir zeigen, dass ungerade periodische Ketten sowie Ketten mit Dirichlet-Randbedingungen eine Birkhoff-Normalform vierter Ordnung zulassen, die zudem in fast allen Fällen Kolmogorovs Nichtentartungsbedingung erfüllt, sodass wir das klassische KAM-Theorem anwenden können, was eine seit langem bestehende Vermutung bestätigt.

Für gerade periodische Ketten erhalten wir eine resonante Normalform vierter Ordnung, von der wir zeigen, dass sie integrierbar ist. Zudem analysieren wir die Blätterung des Phasenraums dieser integrablen Approximation gerader periodischer Ketten in Niveaumengen der Integrale.

Abstract

In this thesis we study the normal form theory of one-dimensional Fermi-Pasta-Ulam chains. We prove that odd periodic chains and chains with Dirichlet boundary conditions admit a fourth order Birkhoff normal form, which satisfies Kolmogorov's nondegeneracy conditions for almost all parameter values. Hence we can apply the classical KAM theorem to these types of chains, thereby proving a long standing conjecture.

For even periodic chains we obtain a resonant fourth order normal form, which we show to be integrable. Furthermore we analyze the foliation of the phase space of this integrable approximation of even periodic chains into level sets of the integrals.

Contents

Introduction **ix**

1 Results **1**
1.1 Setup 1
1.2 Odd periodic chains 3
1.3 Even periodic chains 4
1.4 Dirichlet chains 7
1.5 Geometry of the moment map 8
1.6 Applications 10
1.7 Related work 11

2 Theoretical background **13**
2.1 Hamiltonian systems 13
2.2 Integrable systems 15
2.3 Birkhoff normal form 16
2.4 Perturbed integrable systems 17

3 Normal form computations **21**
3.1 Odd periodic chains 21
3.2 Even periodic chains 31
3.3 Dirichlet chains 35

4 Nondegeneracy and convexity **41**
4.1 Odd periodic chains 41
4.2 Dirichlet chains 48

5 Foliation by the moment map **53**
5.1 Foliation of $\mathcal{P}_0$ 54
5.2 Foliation of $\mathcal{P}_{\frac{N}{4}}$ for $\frac{N}{4} \in \mathbb{Z}$ 54
5.3 Foliation of $\mathcal{P}_k$ for $0 < k < \frac{N}{4}$ 57

6 Discussion and Outlook **71**

A Details of section 3.1 **75**
A.1 Relative coordinates 75
A.2 Birkhoff normal fom up to order two 76

B Nonresonance lemma **79**

C Proof of Lemma 3.3.1 **87**

D Proof of Lemma 4.1.5 **91**

E Spectrum of the matrix P^D **95**

F Critical points of $\mathcal{M}_\gamma$ of rank 0 **99**

Acknowledgements **111**

Curriculum Vitae **113**

List of Figures

5.1 Subsets of parameters (γ, r) with hyperbolic dynamics of X_γ for $N = 100$ and $k = 1, 5, 10, 24$ 60
5.2 Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 1$, $\gamma = -1.35$ . . 65
5.3 Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 1$, $\gamma = -0.5$ . . 66
5.4 Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 1$, $\gamma = 0.35$ . . . 66
5.5 Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 0.3$, $\gamma = -2.5$. 67
5.6 Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 0.3$, $\gamma = -1$ . . 67
5.7 Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 0.3$, $\gamma = 6$ . . . 67

Introduction

And certainly the atoms did not move by volition,
nor did they place themselves by sharp intelligence,
nor did they agree what movements to produce,
but they, being many and moving about in many ways,
are constantly being buffeted and given motion,
and by trying every kind of combination
and motion, finally they fall into the arrangements
and the patterns of which the sum of things consists.

Lucretius [50] (Book I, lines 1021-1028)

The Fermi-Pasta-Ulam problem

The simulations performed in the early 1950's by the physicist Enrico Fermi (1901-1954), the computer scientist John Pasta (1918-1981), and the mathematician Stanislaw Ulam (1909-1984), in a "professional configuration foreshadowing the disciplinary alliance of the future" [84], can without exaggeration be considered as having revolutionized the field of dynamical systems, which for about half a century had seen less progress than other areas of mathematical physics such as quantum mechanics and general relativity.

Before discussing the simulations of Fermi, Pasta, and Ulam (in the sequel called "FPU") and their paradoxical outcome in greater detail, let us briefly "set the stage" by giving a quick review of the situation in the field of dynamical systems in the middle of the twentieth century.

The entire history of the theory of mechanics, and in particular dynamical systems, is also a history of mutual influence between mathematical and physical research, a process which is continuing until the present time. The formalization of mechanics at the beginning of the seventeenth century is associated with the names of Galileo Galilei (1564-1642) and Johannes Kepler (1571-1630), who discovered some of the laws of terrestrial and celestial mechanics, respectively. It was Isaac Newton (1643-1727), who showed that the phenomena in these two areas could actually be described by the same principles, for the formulation of which he also (co-)invented calculus. The laws of mechanics were then reformulated by Joseph-Louis Lagrange (1736-1813) and William Rowan Hamilton

(1805-1865), who formulated them in terms of evolutions in configuration and phase space, respectively, and it is the framework of Hamilton which is still widely used to describe (classical) mechanical systems, but which also turned out to be a suitable starting point for the description of quantum-mechanical systems. An important contribution was Joseph Liouville's (1809-1882) theorem stating that if the energy of a system is conserved, then any volume of initial conditions in phase space must be conserved throughout the evolution. Towards the end of the nineteenth century, the investigation of physical systems with a (very) large number of particles led to the new field of statistical mechanics (pioneered, among others, by Ludwig Boltzmann (1844-1906)), which soon was (and still is) interwoven with probability theory, since the macroscopic description of such systems primarily consists of probabilistic statements.

On the other hand, it was Henri Poincaré (1854-1912) who pioneered the *qualitative* study of dynamical systems, since it soon became clear that many important systems could not be solved analytically. Poincaré also developed the method of perturbations to investigate systems which could be considered as small perturbations of an integrable system, and it is precisely this method that Kolmogorov used half a century later in his seminal work. And even though Kolmogorov applied this method to systems which are not a priori described statistically, the main result of his theory is a statement of somewhat probabilistic nature - we will formulate it precisely in this thesis.

As mentioned above, in the first half of the twentieth century, despite some important contributions e.g. by George David Birkhoff (1884-1944), the theory of dynamical systems did not evolve as rapidly as other areas of mathematical physics. A lot of work was devoted to the notion of *ergodicity*, and most people believed in the "ergodic hypothesis", namely that arbitrarily small perturbations could turn an integrable system into an ergodic one (on each energy surface). Ironically, it was Fermi [19] himself who published a "proof" of this hypothesis, which however later turned out to be incorrect.

When the MANIAC-I computer was built in 1952 by the Theoretical Division (headed by Nicholas Metropolis (1915-1999)) of the Los Alamos National Laboratory, it was Fermi's ingenious idea to use it also as a tool for the simulation of physical systems. Thus he proposed to test the ergodicity hypothesis mentioned above on a comparatively simple system. Precisely, he intended to observe energy sharing among *nonlinearly* coupled rigid masses in one-dimensional chains with fixed endpoints (in the sequel called FPU chains). However, instead of the expected outcome, the results would eventually turn out to be "a challenge for the foundation of physics" [13]. Let us first briefly turn to a system of *linearly* coupled masses.

The behavior of such a system, i.e. a system in which the force on each mass point depends linearly on the distance between itself and its two nearest neighbors, is completely predictable and best described in terms of normal or Fourier modes, in which the Hamiltonian (the total energy) of the system is the sum of the energies of the single modes, i.e. the system can be described as a system of *uncoupled* harmonic oscillators with no exchange of energy between different modes. When the forces between the masses are assumed to be nonlinear, as in the setup of FPU's simulation, additional coupling terms appear in the Hamiltonian, which according to the principles of statistical mechanics led Fermi to the expectation that the energy would eventually be equally distributed among the different modes (equipartition), or, as Weissert [84] puts it, "energy should march through the sequence of harmonic modes like champagne spilling down a pyramid of glasses". It was the intent of the FPU experiments to measure the rate of this expected "thermalization". The number of particles was chosen as 16, 32, or 64 (apparently related to the binary arithmetic of the computer used for the simulation), and the simulations were performed over 14'000 to 19'000 time cycles.

However, this thermalization was *not* observed, the energy was not equipartitioned among the Fourier modes. As initial condition, the entire energy was concentrated in the first mode, and it never seemed to be dispersed beyond the first few modes. More precisely, the energy seemed to oscillate between these first few modes in a "quasi-periodic" way. As stated in [16], the FPU "paradox" (as the results were henceforth called) "shows that nonlinearity is not enough to guarantee equipartition of energy". This apparent contradiction to the expectations demanded an explanation - however, at the time (1954), there was no general theory accomplishing this task, and Fermi immediately recognized the importance of the observations. Unfortunately, due to his death in November 1954, he could not contribute to resolving the paradox any more - his death also considerably delayed the publication of the surprising results. The year in which the report [20] of Fermi, Pasta, and Ulam was finally published (1955) is now widely considered as the "birth" of the FPU problem, and in the last few years numerous reviews on the history of the FPU problem have been published (see e.g. [23, 12, 7, 88]), especially at the occasion of the fiftieth centenary (2005) of the publication of the original report.

Since then, various explanations have been offered to explain the surprising results of the FPU simulations. Following the review articles [7] or [16], one can distinguish two types of explanations, the perturbative and the "soliton-based" approaches. In this thesis, we primarily follow the former approach and show that for all three types of chains and all parameter values, the FPU Hamiltonian can be approximated by an integrable system up to fourth order. In other words, *we construct a fourth-order integrable model for FPU chains.* Furthermore, in the case of the odd periodic and the Dirichlet chains, we show that for almost all parameter values, the KAM theorem can be applied locally around the equilibrium point. The KAM theorem is a result by Andrej Kolmogorov (1903-1987), Vladimir Arnol'd (*1937), and Jürgen Moser (1928-1999) asserting that the "majority" of the orbits of a slightly perturbed integrable system remains quasi-periodic, under a certain nondegeneracy condition on the frequencies of the unperturbed system. We will return to it in a moment and precisely state it in chapter 2.

Even though we thus rigorously confirm the long-standing conjecture that the KAM theorem can be applied to FPU chains, it seems unlikely that this already "explains" the FPU paradox, since it remains unclear whether the energy levels and initial conditions chosen by Fermi, Pasta, and Ulam fit into the "scheme" of the KAM theorem. In particular, since the admissible energy levels for our application of KAM to FPU appear to be becoming smaller and smaller as the number of particles tends to infinity, it seems rather unlikely that the KAM theorem is sufficient for the desired explanation of the FPU paradox.

However, we do not only rely on the KAM theorem, we also plan to numerically implement the dynamics of our integrable fourth-order approximation. If such an implementation would produce results close to FPU's original results, we think that this would be a considerable contribution towards an explanation of the FPU paradox.

Perturbative approaches were however already proposed before the KAM theorem became well-known among mathematical physicists. The first analytical approach to the FPU problem was given by Ford [21] in 1961, arguing that the missing ergodicity in the FPU was based on arithmetical properties of the unperturbed chain (here, contrary to the approach to be developed in this thesis, "unperturbed" means the system of uncoupled harmonic oscillators). Further work in this direction was done by Jackson [39, 40] and Ford and Waters [22].

The KAM theorem then provided new and strong theoretical support to the claim that "typical" nonlinear systems exhibit nonergodic behavior. Note that although Kolmogorov's original work dates from 1954 and the proofs of his conjecture by Arnol'd and Moser from 1962 and 1963, it took some time before their work became well known among the phycisists working on the FPU problem. It is interesting to note that although "FPU" and "KAM" started more or less in the same year - 1954 -, it took at least a decade before it was realized that the latter possibly could contribute to the explanation of the former. This delay was probably caused or at least prolonged by political reasons (since "FPU" originally was a primarily American and "KAM" a primarily Soviet research area).

The first connection between FPU and KAM was made by Izrailev and Chirikov [38] in 1966. However, in this paper it was not rigorously proved that the FPU system actually fulfills the hypotheses of the KAM theorem, the discussion was more about the admissible relative size of the perturbation beyond which the stability asserted by the KAM theorem would "break down", leading to what today is called "strong stochasticity threshold" (some results in this direction can e.g. be found in [89, 61]). The KAM theorem then became widely known in the physics community through the article by Walker and Ford [83] where it was primarily discussed as a possible explanation of the results of the Hénon-Heiles simulation [31], a connection first observed by Gustavson [29]. Whereas after the discovery of the integrability of the Toda lattice [18, 30, 51], it was clear that the three-particle Hénon-Heiles-system could be treated as a perturbation of the (three-particle) Toda lattice, this remained unclear in the case of FPU chains (with an arbitrary number of particles).

In the 1960's, the Japanese school around Saito, Hirooka, and Toda also became active in the research on the dynamics of nonlinear chains, in the beginning however completely unaware of the (at the time still largely American) discussion of the FPU results. These Japanese researchers even started their own numerical simulations, and according to [16] obtained results which somewhat resembled the FPU results, but they only published them after having finally learned about the FPU experiments.

In the subsequent years, the only significantly new approach towards the application of the KAM theorem to FPU chains was Nishida's idea [58] of using Birkhoff normal forms to obtain a nondegenerate integrable system, which was further elaborated by Sanders [75]. The transformation to these Birkhoff normal forms however requires the validity of certain nonresonance conditions, which Nishida did not prove. It was only recently that Rink [70] proved Nishida's conjecture, i.e. actually carried through the transformations rigorously in certain special cases of the parameter values. In the general case, this has not been fully accomplished yet - as Weissert notes in his book [84] on (the first 20 years of) the history of the FPU problem, "Once again, although the claim was made for KAM as the probable explanation for FPU, the conditions for the theorem had not been established rigorously". As already mentioned above, it is one of the main goals of this thesis to rigorously establish this connection, but with the somewhat broader goal of obtaining an integrable fourth-order approximation to the FPU Hamiltonian.

But first let us mention another theorem providing a stability result for perturbed integrable systems, which is much less well known than the KAM theorem, namely the results of "Nekhoroshev type" asserting stability of the motion of the perturbed system under a condition slightly stronger (and more difficult to check) than Kolmogorov's nondegeneracy, namely "steepness" of the unperturbed Hamiltonian. However, the results of this type hold for *all* sufficiently small perturbations, not just for a *majority* as in the case of the KAM theorem. Thus, this type of results is of "deterministic" rather than "probabilistic" nature, whereas there is the drawback of the slightly stronger assumptions on the unperturbed Hamiltonian. We will give a brief overview of Nekhoroshev's results (and their more recent versions) in section 2.

Another issue to be mentioned in this connection, is the work of the "Italian school" on the concept of *metastability*. It can be considered as a refinement of the research thread initiated by Izrailev and Chirikov connecting FPU chains and the KAM theorem. This concept was first introduced in [24] and has been further developed in the sequel. Its main point is the observation of numerical evidence of the existence of two different time scales. As in FPU's original experiments, as initial condition the entire energy is concentrated in the lowest frequency mode. After a first time scale, one observes (up to an exponentially small tail) a constant energy distribution among the first few low frequency modes (on different energy levels), whereas complete equidistribution of energy among all modes is observed only on a second, much longer time scale. However, a thorough theoretical justification in particular of this second time scale has apparently not yet been obtained.

The other approach towards explaining the FPU results consists of considering the continuum limit of the FPU chain and trying to gain insight into the discretized chain through an investigation of this continuum limit. It was Zabusky [85, 86] who first took this approach in 1962 and 1963, and then together with Kruskal in the famous paper [87] of 1965, where the discovery of "soliton" solutions of the periodic KdV equation (introduced in 1895 by Korteweg and de Vries [44]) was reported. By this they meant solitary wave solutions which have the property of passing through one-another and afterwards almost recovering their initial state despite the nonlinear interaction. Thus, they found a behavior which closely resembled the FPU observations. Even though the relation between the continuous and the discrete models were not made adequately clear at the time, the discovery of solitons subsequently led to a variety of related results. In particular, the periodic KdV equation was shown by Gardner et al. [26, 27] to be completely integrable, which together with the integrability of the Toda lattice mentioned above strongly suggested that this might explain the FPU results. However, here the problem is that even though the KdV equation can be seen as the continuum limit of the FPU and Toda chains (in a sense which we will not explain precisely, since we do not pursue this approach), it seems to be difficult to explain how the properties of the continuum limit can be used to explain the behavior of the original discrete FPU chain. Until now, there is a lot of research going on in this area, which we cannot exhaustively discuss. Moreover, the FPU problem also has connections with many other areas of physics such as e.g. Bose-Einstein condensation and quantum chaos - we refer to [7] for a review of some of these issues. Furthermore, there is a huge amount of research on the Toda lattice, a special FPU chain whose especially strong integrability properties lead to many important conclusions. In particular, via the "Lax formalism" [47] one can precisely consider the Toda lattice as a discrete analogue of the KdV equation, and it is one of our ongoing research projects [32, 33, 34] to develop an analogue of Kappeler and Pöschel's normal form theory for the KdV equation [42] for the periodic Toda lattice.

Concerning the Toda lattice, we emphasize that even though it is not the subject of this thesis, it has served as a motivation for our work on arbitrary FPU chains. In particular, it seems to be very important that the family of dynamical systems given by the family of FPU chains contains an integrable system, namely the Toda lattice! From a methodological point of view, this seems to be the "bottom line" of our work.

Of course, the two main threads of explanation of the FPU results, the perturbative approach and the observation of solitons in the continuum limit, do not contradict each other - as Weissert [84] notes, "it might also be said that each of these solutions describes essentially the same phenomenon from a slightly different perspective". In any case, the whole history of the FPU problem can be seen as another excellent example of the mutual influence of physical and mathematical research - in this case made possible by the availability of digital computers for scientific research. The research areas "FPU" and "KAM" both started around 1954, but as already noted, it took more than 10 years before it was realized that the latter could contribute to the explanation of the former.

Moreover, FPU chains are one of the most prominent examples of systems outside the realm of celestial mechanics to which the KAM theorem has been proved to be applicable. This remark holds even more in the case of the Nekhoroshev theorem - here we however do not fully investigate for which parameter values it can be applied, since Nekhoroshev's original criteria of "steepness" are quite difficult to check. We only apply a recent version of the theorem proven by Pöschel where the steepness is replaced by the stronger notion of convexity (but which also holds around an elliptic equilibrium and not only around a torus of full dimension).

Besides contributing to the discussion of the FPU paradox, we also consider this thesis to be a "case study" in the theory of normal forms and perturbations of integrable systems, in the sense that we have applied the theoretical tools of these areas to a specific system (or a specific class of systems) which has played an important role in the development of dynamics in general. Finally, in the case of even periodic chains, where it turns out that we cannot directly apply the KAM theorem due to resonances, our analysis of the geometry of the moment map of the corresponding integrable system reveals surprisingly rich dynamics.

To conclude this introduction, let us make some remarks concerning the epistemological significance of the research on FPU chains. As Weissert [84] notes, the fact that the obtained results contradicted the assumptions is not the only issue to be considered: "In the history and philosopy of science, the general problem of experimental evidence that contradicts the hypothesis underlying the experiment itself is not new. However, the FPU problem is the first such case where the evidence came from the results of a simulation instead of an experiment."

As remarked in [7], the FPU problem can actually be seen as having initiated a new method of research in the physical sciences besides theoretical and experimental physics, namely a "synergetic" cooperation between phycisicts and computers, a term already used by Ulam [80]. The abundant use of this approach nowadays (also in many areas outside physics) easily lets us forget that 50 years ago, the use of computers not just as a simple calculational device, but as a tool for studying "entire" physical systems, was revolutionary. The main idea of this type of experiments consists of letting theoretical predictions and numerical studies mutually influence each other, in particular of unexpected numerical results giving rise to new theoretical insights.

The use of computers as a tool for the simulation of physical processes itself raises a number of epistemological issues. The main one probably is the question of why one should believe that results of a simulation actually tell us something about physical reality. Apart from the problems arising from the discretization of continuous processes necessary for a numerical implementation, one always makes some approximations and simplifying assumptions in the formulation of specific mathematical models. And there are also the philosophical issues of whether experimenters are biased in the interpretation of the outcomes of their results by certain implicit assumptions (regardless of whether the experiments are performed in the laboratory or on a computer). Since it is far beyond the scope of this thesis to discuss these issues, we just mention some literature where these questions are discussed, e.g. [82] or [25]. In any case, we are convinced that the FPU experiments and all the work stimulated by them are an extremely interesting "case study" for the philosophy and history of science. For instance, it would be very interesting to investigate in which sense the research on the

FPU problem and its related issues can be seen as a "paradigm change" in the sense of Kuhn [45] or how Kuhn's notions should be further developed in order to cope with the doubtless "revolutionary" research initiated by the paradoxical FPU results. We will not discuss these questions in this thesis.

Summarizing our results, we prove that FPU chains can be seen as higher order perturbations of a fourth order integrable system, and that in the case of odd periodic and Dirichlet chains the classical KAM theorem can be applied to these chains locally around the fixed point, i.e. for low energies, for almost all parameter values. For the even periodic chains, we cannot apply the classical KAM thoerem, but we investigate the dynamics of the moment map of the corresponding integrable system, thereby finding hyperbolic or elliptic dynamics, depending on suitably chosen bifurcation parameters.

Outline of the thesis

Let us first remark that this thesis essentially is an extended version of our papers [35] and [36]. In chapter 1 we begin by presenting the formal setup of the FPU model, thereby emphasizing several special cases of certain parameter values which are of particular interest. By distinguishing between different types of parity and boundary conditions, we arrive at three different types of chains, the odd periodic, the even periodic, and the Dirichlet chains. We then present all our results on these three types of chains. Whereas for the odd periodic and the Dirichlet chains we obtain Birkhoff normal forms up to order four and some nondegeneracy and convexity results, allowing us to apply the perturbation theory results by KAM and Nekhoroshev, for the even periodic chain we obtain a *resonant* normal form up to order four. We show that this (truncated) resonant fourth order normal form is a completely integrable system and analyze the foliation of its phase space by the moment map given by its integrals. In particular, we show that this integrable system exhibits hyperbolic dynamics.

In chapter 2, we give an overview of the theoretical background of our work. We first review the notions of a Hamiltonian system and the special case of an integrable one, and then explain what we mean by a Birkhoff normal form up to a certain order. Afterwards we discuss the KAM and Nekhoroshev theorems, and we mention some recent improvements of these theorems which have increased their applicability by weakening their hypotheses. In particular, we discuss the different types of nondegeneracy and convexity properties of the Hessian of the unperturbed Hamiltonian which are necessary for the application of these theorems.

The following chapter 3 is the central part of the thesis. Here we perform all computations necessary for the proof of the normal form theorems on our three types of chains. These computations essentially consist of carrying through a series of transformations bringing the FPU Hamiltonian into the desired form. Even though these computations may seem unmotivated, we try to convince the reader that they are strongly inspired by our work on the periodic Toda lattice (which is not contained in this thesis). Nevertheless, all computations are completely self-contained. We also try to emphasize the crucial role played by the parity of the number of particles, as in the case of an even number of particles in the periodic chain there are certain fourth order resonances which do not appear in the odd case. For Dirichlet chains, we do not have to repeat all calculations of the periodic chain - it turns out that these chains can be treated as an invariant submanifold of even periodic chains, and in this special case of even periodic chains, the fourth order resonances mentioned above are no obstruction to the transformation to Birkhoff normal form of order four. Finally, in the general case (of even periodic chains) we prove the integrability of the truncated fourth order resonant normal form, which comes somewhat surprisingly, since it was previously assumed that we have this property only in certain special cases.

Chapter 4 contains the proofs of the theorems on the nondegeneracy and convexity properties of the Hessian of the odd periodic and Dirichlet chains. We do not only prove that the Hessian of these two chains is nondegenerate for almost all parameter values (in a sense to be made precise), we also derive some explicit formulas and asymptotic estimates for some of the exceptional parameter values, i.e. those where the Hessian is not nondegenerate. For some of these exceptional parameter values, we also investigate whether the alternative notion of isoenergetic nondegeneracy holds.

In chapter 5, the third and last major part of the thesis, we study the geometry of the phase space of the truncated resonant normal form of the even periodic chain. We investigate how the moment map given by the integrals found in chapter 3 foliate this phase space into invariant level sets, and it turns out that one can find a very rich geometry. Distinguishing between regular and critical points of the phase space, we perform various reductions depending on the rank of the differential of the moment map at these critical points in order to gain further insight into the structure of the level sets associated to these critical points. In particular, after reducing to two degrees of freedom, we find four critical points of rank zero, and we obtain a bifurcation in the space of suitably chosen parameters which determine the type (elliptic or hyperbolic) of these critical points, and we discuss the question of homo- or heteroclinic orbits. After another reduction, we briefly discuss the remaining critical points.

Finally, in chapter 6 we discuss the relevance of our results and possible directions for future research. The question of relevance is not easy to be answered and in particular depends on the results of some ongoing numerical work implementing some of our transformations.

Almost all proofs of the theorems presented in this thesis contain a lot of very explicit and tedious computations. Therefore and in order not to destroy the “direct line of arguments”, we put some of these computations in appendices. Nevertheless, we point out that these appendices are an essential part of the thesis. Appendix A contains some computations in the transformation of the odd chain into Birkhoff normal form up to order two, in particular the (very elementary) reduction to relative coordinates, since we are not interested in the (linear) motion of the center of mass coordinate. Appendix B contains a complete classification of the fourth order resonances - it is essentially an extended and more detailed version of number theoretic results of the literature. However, it may be possible to prove (some of) these results using the integrability of the (full) periodic Toda lattice - as a proof of a number theory result by dynamical systems methods this would also be interesting from a methodological viewpoint. In Appendix C we give the details of the treatment of the Dirichlet chain as an invariant submanifold of the even periodic chain. Appendix D contains the proof of a combinatoral lemma necessary to prove one of the results on the asymptotic width of the "convexity interval" of the odd periodic chain. Whereas this result may be of questionable independent interest, we have included it, together with its lengthy proof, because we consider the combinatorial method of its proof to be of some interest. In Appendix E we discuss the spectral properties of a matrix which is important for the non-degeneracy discussion of the Dirichlet chain. Finally, Appendix F contains the (straightforward) classification of the four fixed points of rank zero of the reduced moment map of the truncated resonant normal form of the even periodic chain mentioned above.

Chapter 1

Results

In this first chapter we state all our results on normal forms and nondegeneracy properties of one-dimensional FPU chains. First we describe the setup, and then we list our results and their applications, separated by the type of boundary conditions and the parity of the number of particles. Whereas the original FPU simulations were performed for Dirichlet boundary conditions, for the theoretical treatment it turns out to be convenient to first investigate chains with periodic boundary conditions and then treat chains with Dirichlet boundary conditions as an application of the former ones.

1.1 Setup

As explained in the introduction, in this thesis we consider nonlinear chains with N particles of equal mass, normalized to be one, as introduced by Fermi, Pasta, and Ulam, and known as FPU chains. Let us now give the precise model of such chains. A FPU chain consists of a string of particles moving on the line or the circle interacting only with their nearest neighbors through nonlinear springs. Its Hamiltonian is given by

$$H_V = \frac{1}{2}\sum_{n=1}^{N} p_n^2 + \sum_{n=1}^{N} V(q_n - q_{n+1}), \tag{1.1}$$

where $V : \mathbb{R} \to \mathbb{R}$ is a smooth potential. The corresponding Hamiltonian equations read ($1 \leq n \leq N$)

$$\begin{aligned} \dot{q}_n &= \partial_{p_n} H_V = p_n, \\ \dot{p}_n &= -\partial_{q_n} H_V = -V'(q_n - q_{n+1}) + V'(q_{n-1} - q_n). \end{aligned}$$

Here q_n denotes the displacement of the n'th particle from its equilibrium position and p_n its momentum. We assume either N particles and periodic boundary conditions,

$$(q_{i+N}, p_{i+N}) = (q_i, p_i) \quad \forall i \in \{0, 1\}, \tag{1.2}$$

or N' (moving) particles and Dirichlet boundary conditions, i.e. fixed endpoints,

$$q_0 = q_{N'+1} = 0. \tag{1.3}$$

We will treat Dirichlet chains as an application of periodic chains with an even number of particles, hence we first concentrate on periodic chains. Although Dirichlet chains can thus be viewed "structurally" as an application of the theory of periodic chains, they are of great independent interest, in particular because the FPU simulations were originally carried through for chains with Dirichlet boundary conditions.

Without loss of generality, the potential $V : \mathbb{R} \to \mathbb{R}$ is assumed to have a Taylor expansion at 0 of the form

$$V(x) = \kappa \left(\frac{1}{2}x^2 - \frac{\alpha}{3!}x^3 + \frac{\beta}{4!}x^4 + \dots \right), \tag{1.4}$$

where κ is the (linear) spring constant normalized to be 1 and $\alpha, \beta \in \mathbb{R}$ are parameters measuring the strength of the nonlinear interaction. The minus sign in front of the parameter α in the expansion (1.4) turns out to be convenient for later computations. Substituting the expression (1.4) for V into (1.1), the corresponding expansion of H_V is given by

$$H_V = \frac{1}{2}\sum_{n=1}^{N} p_n^2 + \frac{1}{2}\sum_{n=1}^{N}(q_{n+1}-q_n)^2 + \frac{\alpha}{3!}\sum_{n=1}^{N}(q_{n+1}-q_n)^3 + \frac{\beta}{4!}\sum_{n=1}^{N}(q_{n+1}-q_n)^4 + \dots . \tag{1.5}$$

For any FPU chain, the total momentum $P = \frac{1}{N}\sum_{n=1}^{N} p_n$ is an integral of motion, and therefore the center of mass $Q = \frac{1}{N}\sum_{n=1}^{N} q_n$ evolves with constant velocity. Hence any FPU chain can be viewed as a family of Hamiltonian systems of $N-1$ degrees of freedom, parametrized by the vector of initial conditions $(Q, P) \in \mathbb{R}^2$ with Hamiltonian independent of Q. In particular, for $N = 2$ any FPU chain is integrable, and hence we will concentrate on the case $N \geq 3$. Further note that for any vector $(Q, P) \in \mathbb{R}^2$, the origin in $\mathbb{R}^{2N-2}$ is an equilibrium point of the corresponding system. The momentum of such an equilibrium point is given by the constant vector $(p_1, \dots, p_N) = P\,(1, \dots, 1)$.

The frequencies $(\omega_k^0)_{1 \leq k \leq N-1}$ of the linearization of an arbitrary FPU chain at $(q, p) = (0, 0)$ are easily computed to be

$$\omega_k^0 = 2 \sin \frac{k\pi}{N}.$$

The corresponding resonance lattice is given by

$$\left\{ l = (l_1, \dots, l_{N-1}) \in \mathbb{Z}^{N-1} \,\middle|\, \sum_{k=1}^{N-1} l_k \sin \frac{k\pi}{N} = 0 \right\}$$

and generated by the vectors $l^{(k)}$, $1 \leq k \leq N-1$, defined by $l^{(k)} = e_k - e_{N-k}$, where e_i, $1 \leq i \leq N-1$, denotes the standard basis in $\mathbb{R}^{N-1}$.

It turns out that the properties of periodic chains near the equilibrium point strongly depend on the parity of the number N of particles. If N is odd, these chains can be transformed into Birkhoff normal form of order four, whereas if N is even, there are resonances making the analogous transformations impossible. We do not present a group-theoretic explanation of this fact. On the other hand, in the case of Dirichlet boundary conditions, the chains with an odd and an even number N' of particles behave similarly, i.e. these chains can be transformed into Birkhoff normal form of order four, independently of the parity of N'.

There are some parameter values in the expansion (1.5) which are of special importance, because of historical, phenomenological, or structural reasons. The case $\alpha = 0$ is known as the *β-chain*, the case $\beta = 0$ as the *α-chain*, and the case $\beta = \alpha^2$ is a fourth order approximation of the *Toda lattice*. The potential of the full Toda lattice is the exponential function, i.e. $V(x) = \kappa e^{-x}$, introduced by Toda [79] and extensively studied in the sequel - for an overview see e.g. [78]. It turns out that the (full) Toda lattice with periodic boundary conditions is completely integrable, as was shown independently by Flaschka [18], Hénon [30], and Manakov [51], and as mentioned in the introduction, it is one of our ongoing projects [32, 33, 34] to develop a normal form theory for the periodic Toda lattice in analogy to Kappeler and Pöschel's work on the periodic KdV equation [42]. In the sequel, we will however denote by Toda chain any FPU chain with $\beta = \alpha^2$.

1.2 Odd periodic chains

For any point $(x, y) = (x_k, y_k)_{1 \leq k \leq N-1} \in \mathbb{R}^{2N-2}$ we introduce the variables $I = (I_k)_{1 \leq k \leq N-1} \in \mathbb{R}^{N-1}$ by

$$I_k = \frac{1}{2}(x_k^2 + y_k^2). \tag{1.6}$$

Further we define the function $H_{\alpha,\beta} : \mathbb{R}^{N-1} \to \mathbb{R}$, given by

$$H_{\alpha,\beta}(I) := 2\sum_{k=1}^{N-1} \sin\frac{k\pi}{N} I_k + \frac{1}{4N}\sum_{k=1}^{N-1} c_k I_k^2 + \frac{\beta-\alpha^2}{2N} \sum_{\substack{l \neq m \\ 1 \leq l,m \leq N-1}} \sin\frac{l\pi}{N} \sin\frac{m\pi}{N} I_l I_m, \tag{1.7}$$

where $c_k \equiv c_k(\alpha, \beta) := \alpha^2 + (\beta - \alpha^2)\sin^2\frac{k\pi}{N}$.

Our main results on odd periodic chains are the following ones:

Theorem 1.2.1. *Let $\alpha, \beta \in \mathbb{R}$ with $(\alpha, \beta) \neq (0, 0)$. If $N \geq 3$ is odd, then any periodic FPU chain admits a Birkhoff normal form of order 4. More precisely, there are canonical coordinates $(x_k, y_k)_{1 \leq k \leq N-1}$ so that the Hamiltonian of* any *FPU chain, when expressed in these coordinates, takes the form*

$$\frac{NP^2}{2} + H_{\alpha,\beta}(I) + O(|(x, y)|^5)$$

with $H_{\alpha,\beta}(I)$ *given by (1.7).*

Corollary 1.2.2. *Near the equilibrium state, any FPU chain with an odd number* N *of particles can be approximated up to order* 4 *relative to its center of mass coordinates by an integrable system of* $N-1$ *harmonic oscillators which are* coupled *at fourth order except if* $\beta = \alpha^2$ *(Toda case).*

Denote by $Q_{\alpha,\beta}$ the Hessian of $H_{\alpha,\beta}(I)$ at $I = 0$. Note that $Q_{\alpha,\beta}$ is an $(N-1) \times (N-1)$ matrix which only depends on the parameters α and β.

Theorem 1.2.3. *(i) For any given* $\alpha \in \mathbb{R} \setminus \{0\}$, $\det(Q_{\alpha,\beta})$ *is a polynomial in* β *of degree* $N-1$ *and has* $N-1$ *pairwise different real zeroes. When listed in increasing order, the zeroes* $\beta_k = \beta_k(\alpha)$ *satisfy*

$$0 < \beta_1 < \alpha^2, \quad 2\alpha^2 < \beta_2 < \ldots < \beta_{N-1}$$

and contain the $\frac{N-1}{2}$ *distinct numbers*

$$\alpha^2 \cdot \left(1 + \sin^{-2} \frac{k\pi}{N}\right) \quad \left(1 \le k \le \frac{N-1}{2}\right).$$

For these $\frac{N-1}{2}$ *numbers,* $Q_{\alpha,\beta}$ *is* not *isoenergetically nondegenerate.*

When considered as functions $\beta_k = \beta_k^{(N)}(\alpha)$ *of* N*, the zeroes* β_1 *and* β_2 *satisfy*

$$\beta_1 \to \alpha^2, \quad \beta_2 \to 2\alpha^2 \qquad (N \to \infty). \tag{1.8}$$

Moreover index$(Q_{\alpha,\beta})$*, defined as the number of negative eigenvalues of* $Q_{\alpha,\beta}$*, is given by*

$$\mathit{index}(Q_{\alpha,\beta}) = \begin{cases} 1 & \text{for } \beta < \beta_1, \\ 0 & \text{for } \beta_1 < \beta < \beta_2, \\ N-2 & \text{for } \beta > \beta_{N-1}. \end{cases}$$

Hence $H_{\alpha,\beta}$ *is convex if and only if* $\beta_1 < \beta < \beta_2$*.*

(ii) For $\alpha = 0$, $\det(Q_{0,\beta})$ *is a polynomial in* β *of degree* $N-1$*, and* $\beta = 0$ *is the only zero of* $\det(Q_{0,\beta})$*. It has multiplicity* $N-1$*, and the index of* $Q_{0,\beta}$ *is given by*

$$\mathit{index}(Q_{0,\beta}) = \begin{cases} 1 & \text{for } \beta < 0, \\ N-2 & \text{for } \beta > 0. \end{cases}$$

1.3 Even periodic chains

Periodic FPU chains with an even number N of particles do not admit a Birkhoff normal form up to order four due to resonances except if $\beta = \alpha^2$ (Toda case). Applied to even periodic FPU chains, our method of analyzing odd periodic FPU chains leads to a *resonant* Birkhoff normal form up to order four.

Besides the variables $I = (I_k)_{1\le k\le N-1}$ defined by (1.6), it turns out to be convenient to introduce the variables $M = (M_k)_{1\le k\le N-1}$, $J = (J_k)_{1\le k\le N-1}$, and $L = (L_k)_{1\le k\le N-1}$. They are defined on $\mathbb{R}^{2N-2}$ and take values in $\mathbb{R}^{N-1}$, given by

$$M_k = \frac{1}{2}(x_k y_{N-k} - x_{N-k} y_k)\,;\; J_k = \frac{1}{2}(x_k x_{N-k} + y_k y_{N-k})\,;\; L_k = \frac{1}{2}(I_k - I_{N-k})\,. \quad (1.9)$$

Note that for any $1 \le k \le N-1$, $(M_k, J_k, L_k) = (-M_{N-k}, J_{N-k}, -L_{N-k})$ as well as $I_k I_{N-k} = M_k^2 + J_k^2$, or

$$\left(\frac{I_k + I_{N-k}}{2}\right)^2 = M_k^2 + J_k^2 + L_k^2, \quad (1.10)$$

i.e. M_k, J_k, L_k are the Hopf variables expressed in x_k, y_k, x_{N-k}, y_{N-k}. They describe the image of the Hopf map from the three-dimensional sphere of radius $\frac{1}{2}(I_k + I_{N-k})$ centered at the orgin of $\mathbb{R}^4$. Further introduce

$$R_{\alpha,\beta}(J,M) := \frac{\beta-\alpha^2}{4N}\Big(R(J,M) + R_{\frac{N}{4}}(J,M)\Big) \quad (1.11)$$

where

$$R(J,M) = 4 \sum_{1\le k<\frac{N}{4}} \sin\frac{2k\pi}{N}\Big(J_k J_{\frac{N}{2}-k} - M_k M_{\frac{N}{2}-k}\Big) \quad (1.12)$$

and

$$R_{\frac{N}{4}}(J,M) = \begin{cases} J_{\frac{N}{4}}^2 - M_{\frac{N}{4}}^2 & \text{if } \frac{N}{4} \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases} \quad (1.13)$$

Note that for $\alpha, \beta \in \mathbb{R}$ with $\beta = \alpha^2$ (Toda case), the expression $R_{\alpha,\beta}$ vanishes.

Our main results on even periodic FPU chains are the following ones:

Theorem 1.3.1. *Let $\alpha, \beta \in \mathbb{R}$ with $(\alpha,\beta) \ne (0,0)$. If $N \ge 4$ is even, there are canonical coordinates $(x_k, y_k)_{1\le k\le N-1}$ so that the Hamiltonian of* any *FPU chain, when expressed in these coordinates, takes the form $H_V^{trunc}(I,J,M) + O(|(x,y)|^5)$ where*

$$H_V^{trunc}(I,J,M) = \frac{NP^2}{2} + H_{\alpha,\beta}(I) - R_{\alpha,\beta}(J,M) \quad (1.14)$$

and where $H_{\alpha,\beta}(I)$ and $R_{\alpha,\beta}(J,M)$ are given by (1.7) and (1.11), respectively.

As already mentioned, the Hamiltonian H_V in general (i.e. if $\beta \ne \alpha^2$) cannot be transformed into Birkhoff normal form up to order four due to resonances. Nevertheless, the Hamiltonian truncated at order four, H_V^{trunc}, given by (1.14), can be proved to be integrable. The form of the resonance lattice introduced above suggests that $I_k + I_{N-k}$ $(1 \le k \le \frac{N}{2})$ are integrals of H_V^{trunc} in involution.

To find the remaining commuting integrals we express $H_{\alpha,\beta}(I)$ in terms of $I_k + I_{N-k}$ $(1 \le k \le \frac{N}{2})$ and a remainder term,

$$H_{\alpha,\beta}(I) = H^{(2)}(I) + H^{(4)}_{\alpha,\beta}(I) + \frac{1}{2N}\sum_{k=1}^{\frac{N}{2}-1} d_k^- I_k I_{N-k} \tag{1.15}$$

where

$$H^{(2)}(I) = 2\sum_{k=1}^{\frac{N}{2}-1} \sin\frac{k\pi}{N}(I_k + I_{N-k}) + 2I_{\frac{N}{2}},$$

$$\begin{aligned} H^{(4)}_{\alpha,\beta}(I) = &\frac{1}{4N}\sum_{k=1}^{\frac{N}{2}-1} d_k^+ (I_k + I_{N-k})^2 + \frac{\beta}{4N} I^2_{\frac{N}{2}} \\ &+\frac{\beta-\alpha^2}{N} I_{\frac{N}{2}} \sum_{k=1}^{\frac{N}{2}-1} \sin\frac{k\pi}{N}(I_k + I_{N-k}) \\ &+\frac{\beta-\alpha^2}{2N} \sum_{\substack{1\le k,l<\frac{N}{2}\\ k\ne l}} \sin\frac{k\pi}{N}\sin\frac{l\pi}{N}(I_k + I_{N-k})(I_l + I_{N-l}), \end{aligned}$$

and

$$d_k^- \equiv d_k^-(\alpha,\beta) := -\alpha^2 + (\beta-\alpha^2)\sin^2\frac{k\pi}{N}.$$

By (1.10), one has $I_k I_{N-k} = J_k^2 + M_k^2$ for any $1 \le k \le \frac{N}{2} - 1$ so that the remainder term $\frac{1}{2N}\sum_{k=1}^{\frac{N}{2}-1} d_k^- I_k I_{N-k}$ in (1.15) can be written as

$$\frac{1}{2N}\left(\sum_{1\le k<\frac{N}{4}} \left(d_k^-(J_k^2 + M_k^2) + d_{\tilde{k}}^-(J_{\tilde{k}}^2 + M_{\tilde{k}}^2)\right) + d_{\frac{N}{4}}^- \left(J^2_{\frac{N}{4}} + M^2_{\frac{N}{4}}\right)\right),$$

where the latter term is defined to be 0 if $\frac{N}{4} \notin \mathbb{Z}$, and where $\tilde{k} \equiv \tilde{k}(k) = \frac{N}{2} - k$. Combined with the expression (1.11) for $R_{\alpha,\beta}(J,M)$ the Hamiltonian H_V^{trunc} in (1.14) then takes the form

$$H_V^{trunc} = \frac{NP^2}{2} + H^{(2)}(I) + H^{(4)}_{\alpha,\beta}(I) + \frac{1}{2N}\sum_{1\le k\le\frac{N}{4}} K_k(I,J,M) \tag{1.16}$$

where for $1 \le k < \frac{N}{4}$

$$K_k(I,J,M) = d_k^-(J_k^2 + M_k^2) + d_{\tilde{k}}^-(J_{\tilde{k}}^2 + M_{\tilde{k}}^2) - 2(\beta-\alpha^2)\sin\frac{2k\pi}{N}(J_k J_{\tilde{k}} - M_k M_{\tilde{k}}) \tag{1.17}$$

and

$$K_{\frac{N}{4}}(I,J,M) = \begin{cases} -\alpha^2 J^2_{\frac{N}{4}} + (\beta - 2\alpha^2) M^2_{\frac{N}{4}} & \text{if } \frac{N}{4}\in\mathbb{N}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.18}$$

Note that $K_{\frac{N}{4}}(I,J,M)$ can (in the case $\frac{N}{4}\in\mathbb{N}$) also be written as

$$K_{\frac{N}{4}}(I,J,M) = -\frac{\beta-\alpha^2}{2}\left(J^2_{\frac{N}{4}} - M^2_{\frac{N}{4}}\right) + \frac{\beta-3\alpha^2}{2}\left(J^2_{\frac{N}{4}} + M^2_{\frac{N}{4}}\right), \qquad (1.19)$$

in analogy to (1.17).

Theorem 1.3.2. *Let $N \geq 4$ be an even integer. Then for any $\alpha,\beta\in\mathbb{R}$ with $(\alpha,\beta)\neq(0,0)$ the truncated FPU Hamiltonian H_V^{trunc} given by (1.14) is completely integrable. The following $N-1$ quantities are functionally independent integrals in involution:*

$$(I_k + I_{N-k})_{1\leq k\leq\frac{N}{2}},\ (I_k + I_{\frac{N}{2}+k})_{1\leq k<\frac{N}{4}},\ (K_k)_{1\leq k\leq\frac{N}{4}}. \qquad (1.20)$$

Remark: In the case $\beta=\alpha^2$, it follows from (1.17) and (1.19) that the integrals $K_k(I,J,M)$ only depend on the action variables I_k, as $M_j^2+J_j^2 = I_j I_{\frac{N}{2}-j}$ for any $1\leq j<\frac{N}{2}$.

Before turning to the geometry of the moment map of the integrable system of Theorem 1.3.2, we state our normal form and nondegeneracy results for Dirichlet chains.

1.4 Dirichlet chains

As indicated above, FPU chains with Dirichlet boundary conditions can be treated as invariant submanifolds of even periodic chains. To make this precise, consider a chain with N' ($N'\geq 3$, not necessarily even) moving particles and Hamiltonian given by

$$H_V^D = \frac{1}{2}\sum_{n=1}^{N'} p_n^2 + \sum_{n=1}^{N'} V(q_n - q_{n+1}) \qquad (1.21)$$

with boundary conditions (1.3).

Our main results on Dirichlet FPU chains are the following ones:

Theorem 1.4.1. *Let $\alpha,\beta\in\mathbb{R}$ with $(\alpha,\beta)\neq(0,0)$. Then any FPU chain with $N'\geq 3$ moving particles with Dirichlet boundary conditions admits a Birkhoff normal form of order 4, i.e. there are canonical coordinates $(x_k,y_k)_{1\leq k\leq N'}$ so that H_V^D takes the form*

$$\frac{(N'+1)P^2}{2} + H^D_{\alpha,\beta}(I) + O(|(x,y)|^5)$$

where $I=(I_1,\ldots,I_{N'})$ is given by (1.6) and $H^D_{\alpha,\beta}(I)$ is of the form

$$2\sum_{k=1}^{N'} s_k I_k + \frac{1}{16(N'+1)}\sum_{k=1}^{N'}\left(\alpha^2+3(\beta-\alpha^2)s_k^2\right)I_k^2 \quad \underbrace{+\frac{\beta-\alpha^2}{32(N'+1)}I^2_{\frac{N'+1}{2}}}_{\text{only if } \frac{N'+1}{2}\in\mathbb{N}}$$

$$+\frac{\beta-\alpha^2}{16(N'+1)}\left(\sum_{\substack{l\neq m\\ 1\leq l,m\leq N'}} 4s_l s_m I_l I_m - \sum_{k=1}^{N'} s_{2k} I_k I_{N'+1-k}\right), \quad (1.22)$$

where in (1.22), the numbers $s_k := \sin\frac{k\pi}{2N'+2}$ *for any* $1 \leq k \leq N'$ *are pairwise different.*

Corollary 1.4.2. *Near the equilibrium state, any FPU chain with* N' *moving particles and Dirichlet boundary conditions can be approximated up to order* 4 *by an integrable system of* N' *harmonic oscillators which are* coupled *at fourth order except if* $\beta = \alpha^2$ *(Toda case).*

Denote by $Q^D_{\alpha,\beta}$ the Hessian of $H^D_{\alpha,\beta}(I)$ at $I = 0$. Note that $Q^D_{\alpha,\beta}$ is an $N' \times N'$ matrix which only depends on the parameters α and β.

Theorem 1.4.3. *(i) For any given* $\alpha \in \mathbb{R}\setminus\{0\}$, $\det(Q^D_{\alpha,\beta})$ *is a polynomial in* β *of degree* N' *and has* N' *real zeroes (counted with multiplicities). When listed in increasing order, the zeroes* $\beta_k = \beta_k(\alpha)$ *satisfy*

$$\beta_1 \leq \ldots \leq \beta_{\lceil\frac{N'+1}{2}\rceil} < \alpha^2 < \beta_{\lceil\frac{N'+3}{2}\rceil} \leq \ldots \leq \beta_{N'}.$$

Moreover index$(Q^D_{\alpha,\beta})$*, defined as the number of negative eigenvalues of* $Q^D_{\alpha,\beta}$*, is given by*

$$\mathit{index}(Q^D_{\alpha,\beta}) = \begin{cases} \lceil\frac{N'+1}{2}\rceil & \text{for } \beta < \beta_1, \\ 0 & \text{for } \beta_{\lceil\frac{N'+1}{2}\rceil} < \beta < \beta_{\lceil\frac{N'+3}{2}\rceil}, \\ \lfloor\frac{N'-1}{2}\rfloor & \text{for } \beta > \beta_{N'}. \end{cases}$$

Hence $Q^D_{\alpha,\beta}$ *is convex if and only if* $\beta_{\lceil\frac{N'+1}{2}\rceil} < \beta < \beta_{\lceil\frac{N'+3}{2}\rceil}$.

(ii) For $\alpha = 0$, $\det(Q^D_{0,\beta})$ *is a polynomial in* β *of degree* N'*, and* $\beta = 0$ *is the only zero of* $\det(Q^D_{0,\beta})$*. It has multiplicity* N'*, and the index of* $Q^D_{0,\beta}$ *is given by*

$$\mathit{index}(Q^D_{0,\beta}) = \begin{cases} \lceil\frac{N'+1}{2}\rceil & \text{for } \beta < 0, \\ \lfloor\frac{N'-1}{2}\rfloor & \text{for } \beta > 0. \end{cases}$$

1.5 Geometry of the moment map of the truncated even periodic chain

The integrals listed in (1.20) can be partitioned into $\lfloor\frac{N}{4}\rfloor + 1$ groups of integrals which depend only on mutually disjoint subsets of the variables $(x_k, y_k)_{1\leq k\leq N-1}$. As a consequence, the phase space $T^*\mathbb{R}^{N-1}$ of H^{trunc}_V is the direct sum of $\lfloor\frac{N}{4}\rfloor+1$ invariant subspaces, $T^*\mathbb{R}^{N-1} = \bigoplus_{0\leq k\leq\frac{N}{4}} \mathcal{P}_k$, with

$$\mathcal{P}_k = \{(x_k, y_k)_{1\leq k\leq N-1} \in T^*\mathbb{R}^{N-1} | x_j = y_j = 0 \;\forall\, j \notin \{k, N-k, \tilde{k}, N-\tilde{k}\}\},$$

and the foliations of $\mathcal{P}_0$, $\mathcal{P}_{\frac{N}{4}}$ (if $\frac{N}{4} \in \mathbb{N}$), and $\mathcal{P}_k$ for $0 < k < \frac{N}{4}$ into level sets of the integrals can be analyzed separately. We now briefly describe the results obtained by the analysis of the foliations of $\mathcal{P}_{\frac{N}{4}} \cong T^*\mathbb{R}^2$ (if $\frac{N}{4} \in \mathbb{N}$) by the integrals $I_{\frac{N}{4}} + I_{\frac{3N}{4}}$, $K_{\frac{N}{4}}$, and of $\mathcal{P}_k \cong T^*\mathbb{R}^2$ (for $0 < k < \frac{N}{4}$) by the integrals $I_k + I_{N-k}$, $I_{\frac{N}{2}-k} + I_{\frac{N}{2}+k}$, $I_k + I_{\frac{N}{2}+k}$, K_k (whereas $\mathcal{P}_0 \cong T^*\mathbb{R}$ is simply foliated into circles by $I_{\frac{N}{2}}$).

Note that in the case $\beta - \alpha^2 = 0$ the integrals of Theorem 1.3.2 can be replaced by the action variables $I_1, \ldots, I_{N-1}$, so it remains to analyze the case $\beta - \alpha^2 \neq 0$. Instead of α and β, it turns out to be convenient to use the bifurcation parameter

$$\gamma := \frac{\alpha^2}{\alpha^2 - \beta}.$$

The geometry of the moment map $\mathcal{M} \equiv (H, K) : T^*\mathbb{R}^2 \to \mathbb{R}^2$ with $H := I_{\frac{N}{4}} + I_{\frac{3N}{4}}$ and $K := K_{\frac{N}{4}}$ can be analyzed as follows. We use the Hopf variables defined in (1.9), in which K is given by a constant multiple of $(1+\gamma)M^2 + \gamma J^2$ ($M \equiv M_{\frac{N}{4}}$, $J \equiv J_{\frac{N}{4}}$), and observe that the origin of $T^*\mathbb{R}^2$ is the only critical point of $\mathcal{M}$ of rank 0, with $\mathcal{M}^{-1}\{(0,0)\} = \{(0,0)\}$. Then we use a Hopf map to reduce the integral K to level sets of H, obtaining for the reduced vector field X_γ induced by K the formula

$$X_\gamma = \begin{cases} (-2\gamma JL, 2(1+\gamma)ML, -2MJ) & \gamma \notin \{-1, 0\}, \\ (0, L, -J) & \gamma = 0, \\ (-L, 0, M) & \gamma = -1. \end{cases}$$

One then sees that in the case $\gamma \notin \{-1, 0\}$, the reduced vector field X_γ admits six fixed points, four of which are elliptic and the other two hyperbolic fixed points connected by heteroclinic X_γ-orbits.

To analyze the foliation of $T^*\mathbb{R}^4$ by the moment map $\mathcal{M} \equiv (H_1, H_2, G, K) : T^*\mathbb{R}^4 \to \mathbb{R}^4$ with $H_1 := I_k + I_{N-k}$, $H_2 := I_{\frac{N}{2}-k} + I_{\frac{N}{2}+k}$, $G := I_k + I_{\frac{N}{2}+k}$, and $K := K_k$, we proceed similarly. We write 1, 2 for the indices k, $\frac{N}{2} - k$, again introduce the Hopf variables $(M_i, J_i, L_i)_{1 \le i \le 2}$ by (1.9), and, after briefly discussing the critical points of $\mathcal{M}$ of rank one and two, reduce to level sets of H_1 and H_2 through a symplectic reduction given by the product of two Hopf maps. By this procedure, we obtain a reduced moment map

$$\mathcal{M}_\gamma : \mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2} \to \mathbb{R}^2, (M_i, J_i, L_i)_{1 \le i \le 2} \mapsto (G, K_\gamma),$$

where $(r_i)_{1 \le i \le 2}$ are the values of $(H_i)_{1 \le i \le 2}$, together with two reduced Hamiltonian vector fields Y and X_γ of G and K_γ. We show that there exist four critical points of $\mathcal{M}_\gamma$ of rank zero, namely $(M_i, J_i, L_i)_{1 \le i \le 2} = \varepsilon(0, 0, r_1, 0, 0, \pm r_2)$, where $\varepsilon \in \{\pm\}$. It turns out to be convenient to introduce the second bifurcation parameter

$$r := \frac{r_1}{r_2}.$$

Whereas the two points $\varepsilon(0, 0, r_1, 0, 0, -r_2)$ turn out to be elliptic for all parameter values, on the two points $\varepsilon(0, 0, r_1, 0, 0, r_2)$ we prove the following theorem:

Theorem 1.5.1. *Assume that* $1 \le k < \frac{N}{4}$, $0 < r \le 1$, $\varepsilon \in \{\pm\}$, *and* $\gamma \in \mathbb{R}$. *The critical point* $\varepsilon(0,0,r_1,0,0,r_2)$ *of* $\mathcal{M}_\gamma$ *is a* hyperbolic *fixed point of the vector field* X_γ *if and only if*

$$\left|\left(\gamma + \sin^2\frac{k\pi}{N}\right)\sqrt{r} + \left(\gamma + \cos^2\frac{k\pi}{N}\right)\frac{1}{\sqrt{r}}\right| < 2\sin\frac{2k\pi}{N}. \tag{1.23}$$

Otherwise it is an elliptic *fixed point of* X_γ. *If (1.23) is satisfied, the stable and unstable manifolds of* $\varepsilon(0,0,r_1,0,0,r_2)$ *both have dimension two. In the case* $r < 1$, *the connected component of* $\mathcal{M}_\gamma^{-1}\{\varepsilon(r_1 - r_2, 0)\}$ *containing* $\varepsilon(0,0,r_1,0,0,r_2)$ *is a 2-dimensional torus pinched at* $\varepsilon(0,0,r_1,0,0,r_2)$ *and consists of homoclinic* X_γ*-orbits. In the case* $r = 1$, $\mathcal{M}_\gamma^{-1}\{(0,0)\}$ *is a 2-dimensional torus pinched at the two points* $\pm(0,0,r_1,0,0,r_1)$, *and* $\mathcal{M}_\gamma^{-1}\{(0,0)\} \setminus \{\pm(0,0,r_1,0,0,r_1)\}$ *consists of heteroclinic* X_γ*-orbits.*

On the critical points of $\mathcal{M}_\gamma$ with rank $d\mathcal{M}_\gamma = 1$ we have the following result:

Theorem 1.5.2. *Assume that* $1 \le k < \frac{N}{4}$, $0 < r \le 1$, *and* $\gamma \in \mathbb{R}$. *If a point* $(M_i, J_i, L_i)_{1\le i\le 2} \in \mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2} \setminus \{\pm(0,0,r_1,0,0,\pm r_2)\}$ *is a critical point of* $\mathcal{M}_\gamma$ *with rank* $d\mathcal{M}_\gamma = 1$, *then* $(M_2, L_2) = \lambda(M_1, -J_1)$ *for some* $\lambda \in \mathbb{R}$, *and*

$$(r_1^2 - L_1^2)^2 L_2^2 + (r_2^2 - L_2^2)^2 L_1^2 + 2(r_1^2 - L_1^2)(r_2^2 - L_2^2)(2L_1L_2 - (d_{1,\gamma}L_1 + d_{2,\gamma}L_2)^2) = 0.$$

Given any point $(M_1, J_1, L_1) \in \mathbb{S}^2_{r_1} \setminus \{\pm(0,0,r_1)\}$ *there exist at most eight points* $(M_2, J_2, L_2) \in \mathbb{S}^2_{r_2} \setminus \{\pm(0,0,r_2)\}$ *such that* $(M_i, J_i, L_i)_{1\le i\le 2}$ *is a critical point of* $\mathcal{M}_\gamma$ *with rank* $d\mathcal{M}_\gamma = 1$.

By another symplectic reduction, we briefly discuss the type (hyperbolic or elliptic) of these critical points of $\mathcal{M}_\gamma$ with rank $d\mathcal{M}_\gamma = 1$.

1.6 Applications

Although the results desribed in the previous two sections 1.4 and 1.5 are already applications of Theorems 1.3.1 and 1.3.2 on the existence and integrability of the fourth order resonant normal form of even periodic chains, here we discuss those applications of our results which are relevant for an "explanation" of the FPU results.

The first application is that it rigorously follows that the classical KAM theorem can be applied to odd periodic and Dirichlet FPU chains locally around the equilibrium point. As mentioned in the introduction, this confirms long standing conjectures which have never been proved in full generality - we hope to close this gap with this thesis.

Precisely, Theorems 1.2.1 and 1.2.3 allow to apply for any given $\alpha \in \mathbb{R}$ the classical KAM theorem (Theorem 2.4.1) near the equilibrium point to odd periodic chains with Hamiltonian H_V for a real analytic potential $V(x) = \frac{1}{2}x^2 - \frac{\alpha}{3!}x^3 + \frac{\beta}{4!}x^4 + \ldots$ with $\beta \in \mathbb{R} \setminus \{\beta_1(\alpha), \ldots, \beta_{N-1}(\alpha)\}$. Moreover, note that for any given $\alpha \in \mathbb{R} \setminus \{0\}$, the Hessian $Q_{\alpha,\beta}$ of $H_{\alpha,\beta}$ is positive definite for any

β satisfying $\beta_1(\alpha) < \beta < \beta_2(\alpha)$, in particular for $\alpha^2 \leq \beta \leq 2\alpha^2$. Hence one can apply Nekhoroshev's theorem (Theorem 2.4.2) near the equilibrium point to odd periodic chains with Hamiltonian H_V for such β's.

In the case of even periodic chains, we cannot directly apply the classical KAM or Nekhoroshev theorems because of the resonant terms in the fourth order normal form, and it is an open question whether more recent extensions of the KAM theorem (see e.g. [91]) can be applied in this situation.

Moreover, Theorems 1.4.1 and 1.4.3 again allow to apply for any given $\alpha \in \mathbb{R}$ the classical KAM theorem near the equilibrium point to Dirichlet chains for a real analytic potential $V(x) = \frac{1}{2}x^2 - \frac{\alpha}{3!}x^3 + \frac{\beta}{4!}x^4 + \ldots$ with $\beta \in \mathbb{R} \setminus \{\beta_1(\alpha), \ldots, \beta_{N'}(\alpha)\}$. Moreover, as for any given $\alpha \in \mathbb{R} \setminus \{0\}$, $Q^D_{\alpha,\beta}$ is positive definite for $\beta_{\lfloor \frac{N'-1}{2} \rfloor}(\alpha) < \beta < \beta_{\lfloor \frac{N'+1}{2} \rfloor}(\alpha)$, one can apply Nekhoroshev's theorem near the equilibrium point to Dirichlet chains for a potential V with such β's.

As emphasized in the introduction, we do not only rely on the KAM theorem in our attempts towards an explanantion of the FPU results, since it is not clear whether the energy levels and initial conditions are covered by the applications of the KAM and Nekhoroshev theorems to FPU chains just described, especially in the case of large N. However, for all three types of FPU chains we have constructed an approximation of the FPU Hamiltonian up to order four (proven in Theorems 1.2.1, 1.3.1, and 1.4.1) which is completely integrable (which is clear for odd periodic and Dirichlet chains and proven for even periodic chains in Theorem 1.3.2). We plan to numerically implement the dynamics of these integrable approximations and compare the resulting trajectories with the original results of Fermi, Pasta, and Ulam (and more recent simulations of the "full" FPU Hamiltonian), and if these two types of simulations yield "similar" results, we think that our results could be considered as an important step towards explaining the FPU results.

1.7 Related work

The literature on FPU chains is huge, and it seems almost impossible to give a complete overview; we have given some references in the introduction, a lot of citations can also be found in the survey papers [23] and [7].

A large part of our work can be viewed as an extension to the case of arbitrary $\alpha, \beta \in \mathbb{R}$ of work on the β-chain by Rink in various papers [69, 70, 71, 73] and his thesis [72]. In particular, Theorems 1.2.1, 1.2.3, and 1.3.2 are related to results contained in [70], and Theorems 1.4.1 and 1.4.3 to results in [73]. Our study of the foliation of the phase space of the truncated even periodic chain finally is an extension of [71]. It is especially surprising that the integrability of the truncated even periodic chain holds not only for the β-chain, but for the general α-β-chain. Related computations for the periodic β-chain have also been performed by Poggi and Ruffo [66].

Although our work can thus be seen under this viewpoint, it is important to notice that our approach has been shaped not by the consideration of the historically important cases of the α- or the β-chain, but rather by our work on the periodic Toda lattice [32, 33, 34]. (We again emphasize that the proofs of our results are independent of these papers.) It turns out that the same canonical transformations which near the equilibrium bring the Toda lattice into Birkhoff normal form can be used for any FPU chain. Thus, in our view, the most special case of an FPU chain is not the α- or the β-chain, but rather the Toda lattice with its especially strong integrability properties.

It is another question for which potentials $V(x)$ the full FPU Hamiltonian H_V (and not just the truncated Hamiltonian H_V^{trunc}) is integrable. For some contributions in this direction see [11], [68], and [28]; it seems likely that besides the (full) Toda lattice, there are not many other potentials with this property.

One of the most important open problems in the field of FPU chains is the investigation of the dynamics of these chains when the number of particles gets larger and larger - this is strongly related to the “soliton-based” approach mentioned in the introduction. It is likely that our results on FPU chains with Dirichlet boundary conditions can be used for this purpose. For recent contributions in this direction see e.g. [5, 6]. It is actually one of our projects to investigate the “behavior” of our results in the limit $N \to \infty$.

Chapter 2

Theoretical background

In this chapter we explain the theoretical background of the thesis, namely Hamiltonian and in particular integrable systems, Birkhoff normal forms, and some theorems on perturbed integrable systems. We largely follow the exposition in [42].

2.1 Hamiltonian systems

Here we give a brief overview of the abstract Hamiltonian formalism, which is suitable for the mathematical description of physical systems. Let M be a smooth manifold of finite dimension without boundary, which is connected (but not necessarily compact), and let $\mathbb{F} = C^\infty(M)$. (For a review of the notion of a smooth manifold, we refer to any textbook on differential gemometry, e.g. [46].)

Definition 2.1.1. *A* Poisson bracket *on M is a skew-symmetric bilinear map*

$$\{\cdot,\cdot\} : \mathbb{F} \times \mathbb{F} \to \mathbb{F},$$

which satisfies the Leibniz rule*,*

$$\{FG, H\} = F\{G, H\} + G\{F, H\} \quad \forall\, F, G, H \in \mathbb{F},$$

and the Jacobi identity,

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0 \quad \forall\, F, G, H \in \mathbb{F}. \tag{2.1}$$

A smooth manifold with a Poisson bracket is called a Poisson manifold*.*

Furthermore, a flow Φ^t on a Poisson manifold is called a *Poisson system*, if there exists a function $H \in \mathbb{F}$, the *Hamiltonian* of the system, such that

$$\dot{F} := \frac{d}{dt} F \circ \Phi^t|_{t=0} = \{F, H\} \quad \forall\, F \in \mathbb{F}.$$

Since the map $\mathbb{F} \to \mathbb{F}$, $F \mapsto \{F, H\}$ is a derivation, there exists a unique vector field X_H, the *Hamiltonian vector field* associated to H, such that

$$\{G, H\} = X_H G = \langle dG, X_H \rangle \quad \forall\, G \in \mathbb{F},$$

where $\langle \cdot, \cdot \rangle$ denotes the dual pairing between T^*M and TM. Note that due to the skew-symmetry of the Poisson bracket, we also have $\langle dG, X_H \rangle = -\langle dH, X_G \rangle$, hence we can regard X_H as a function of dH, which is linear. In other words, there exists a unique map $K : T^*M \to TM$, the *Poisson structure*, mapping each fiber T_p^*M linearly into T_pM, such that $X_H = KdH$. Since M is finite-dimensional, if the Poisson structure K is *nondegenerate*, i.e. has a trivial kernel, it must be a bijection, and we can consider the inverse $K^{-1} : TM \to T^*M$. This defines a bilinear form ν on vector fields by

$$\nu(X, Y) := \langle K^{-1}X, Y \rangle. \tag{2.2}$$

The form ν defined by (2.2) is skew-symmetric and nondegenerate (since K is a bijection). Moreover, by the Jacobi identity (2.1), $d\nu = 0$, i.e. ν is closed. In other words, ν is a *symplectic form* on M (a closed, nondegenerate 2-form), and (M, ν) is a *symplectic manifold*. Conversely, given a symplectic manifold (M, ν), one defines an isomorphism between TM and T^*M at each point through $X \mapsto \nu \circ X$, the *symplectic structure*, and obtains for any $H \in \mathbb{F}$ the Hamiltonian vector field X_H of H by $X_H = JdH$, which then allows to construct a Poisson bracket on M by

$$\{F, G\} = \nu(X_F, X_G) \quad \forall F, G \in \mathbb{F}.$$

Hence, in our finite-dimensional setting, symplectic forms and nondegenerate Poisson brackets are equivalent notions, and in the sequel we will not strictly distinguish between them.

An important class of diffeomorphisms of a Poisson or symplectic manifold are the diffeomorphisms preserving the underlying structure.

Definition 2.1.2. *A diffeomorphism Φ of a Poisson manifold is called* canonical*, if it preserves the Poisson bracket, i.e. if $\{F, G\} \circ \Phi = \{F \circ \Phi, G \circ \Phi\}$ for any $F, G \in \mathbb{F}$.*

A diffeomorphism Φ of a symplectic manifold is called symplectic*, if it preserves the symplectic form, i.e. if $\Phi^*\nu = \nu$.*

The standard example of a Poisson manifold is $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ with Poisson bracket $\{F, G\}_0 = \langle F_q, G_p \rangle_0 - \langle F_p, G_q \rangle_0$, where $\langle \cdot, \cdot \rangle_0$ denotes the standard scalar product in $\mathbb{R}^n$. For a Hamiltonian H, the coordinate functions evolve by

$$\dot{q}_i = \{q_i, H\}_0 = H_{p_i}, \quad \dot{p}_i = \{p_i, H\}_0 = -H_{q_i}.$$

In terms of symplectic forms, the standard symplectic form obtained from this standard Poisson bracket is $\nu_0 = \sum_{i=1}^n dq_i \wedge dp_i$.

If the position coordinates are identified modulo 2π and thus are angular coordinates, the phase space is $\mathbb{T}^n \times \mathbb{R}^n$, where $\mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$, and the coordinates are denoted by

$$(\theta, I) = (\theta_1, \ldots, \theta_n, I_1, \ldots, I_n)$$

and called *action-angle coordinates*.

2.2 Integrable systems

We call a Hamiltonian system with n degrees of freedom *integrable*, if it admits n functionally independent integrals in involution, which implies that the system can be (formally) solved for any initial data by quadratures.

Definition 2.2.1. *A Hamiltonian system on a Poisson manifold M of dimension $2n$ is called* integrable*, if its Hamiltonian H admits n functionally independent integrals $F_1, \ldots, F_n$ in involution, i.e.*

(i) $\{H, F_i\} = 0$ *for* $1 \leq i \leq n$ *everywhere on* M*,*

(ii) $\{F_i, F_j\} = 0$ *for* $1 \leq i, j \leq n$ *everywhere on* M*, and*

(iii) $dF_1 \wedge \ldots \wedge dF_n \neq 0$ *on an open dense subset of* M*.*

For example, in standard action-angle coordinates $(\theta, I) \in \mathbb{T}^n \times \mathbb{R}^n$ any Hamiltonian which depends only on the action variables $I_1, \ldots, I_n$ is integrable with integrals $F_i = I_i$ for any $1 \leq i \leq n$. This example however is also typical for the case of a general integrable system, at least if one of its leaves (the inverse images $F^{-1}(c)$ for some $c \in \mathbb{R}^n$ and $F = (F_1, \ldots, F_n)$) is compact and connected, as the theorem of Liouville-Arnol'd-Jost-Mineur [4, 41] on the "semi-global" existence of action-angle coordinates and its corollaries assert. We will not cite this theorem here, since we will not apply it in the thesis.

In the sequel, we assume that an integrable Hamiltonian $H = H(I)$ is given in action-angle coordinates $(\theta_i, I_i)_{1 \leq i \leq n}$ with (nonvanishing) integrals $I_1, \ldots, I_n$. The equations of motion are then given by

$$\dot{\theta}_i = \omega_i(I) = \frac{\partial H}{\partial I_i}(I), \quad \dot{I}_i = 0, \quad 1 \leq i \leq n,$$

which can be immediately integrated to $\theta(t) = \theta^0 + \omega(I^0)t$, $I(t) = I^0$. The solution curves are straight lines winding around the underlying invariant torus $T_{I^0} = \mathbb{T}^n \times \{I^0\}$ with constant frequencies

$$\omega(I^0) = (\omega_1(I^0), \ldots, \omega_n(I^0)).$$

These tori are called *Kronecker tori*. The properties of the flow on such a Kronecker torus strongly depend on the arithmetical properties of the frequency vector ω. We distinguish between two cases:

(i) The frequencies $\omega_1, \ldots, \omega_n$ are *nonresonant*, or *rationally independent*, if

$$\langle k, \omega \rangle \neq 0 \quad \text{for all}\, k \in \mathbb{Z}^n \setminus \{0\}.$$

On such a torus, the orbits are dense and the flow ergodic.

(ii) The frequencies ω are *resonant*, or *rationally dependent*, if there exists some $k \in \mathbb{Z}^n \setminus \{0\}$ such that

$$\langle k, \omega \rangle = 0.$$

On such a torus, the orbits are not dense. Each orbit is dense on some lower dimensional torus.

We will see below that it is necessary to single out a sharper version of case (i), the *strongly nonresonant* frequencies (see (2.6)).

2.3 Birkhoff normal form

Another type of integrable systems is given in standard canonical cartesian (rectangular) coordinates $w = (q, p)$ on $\mathbb{R}^n \times \mathbb{R}^n$ with a Hamiltonian of the form

$$H = H(q_1^2 + p_1^2, \dots, q_n^2 + p_n^2).$$

Such systems are integrable with integrals $F_i = q_i^2 + p_i^2$ $(1 \leq i \leq n)$. This type of systems arises in the study of equilibria of Hamiltonian systems. One then also has the variables $I_i = \frac{q_i^2 + p_i^2}{2}$ $(1 \leq i \leq n)$, which can be thought of as action coordinates, but are functions of the rectangular coordinates (q, p).

Consider now such an isolated equilibrium of a Hamiltonian system on some $2n$-dimensional symplectic manifold, i.e. an isolated singular point of the Hamiltonian vector field. Neglecting an irrelevant additive constant, the Hamiltonian, when expressed in the coordinates $w = (q, p)$ near the equilibrium with coordinates $q = 0$, $p = 0$, then has the form

$$H = \frac{1}{2}\langle Aw, w \rangle + \dots$$

where A is the symmetric $2n \times 2n$-Hessian of H at the equilibrium point and the dots stand for terms of higher order in w. We now assume that the equilibrium point $w = 0$ is elliptic, i.e. the spectrum of the linearized system, $\dot{w} = JAw$, is purely imaginary, $\operatorname{spec}(JA) = \{\pm i\lambda_1, \dots, \pm i\lambda_n\}$ with real numbers $\lambda_1, \dots, \lambda_n$. Here $J = \begin{pmatrix} 0 & Id_n \\ -Id_n & 0 \end{pmatrix}$ is the standard symplectic structure of $\mathbb{R}^{2n}$. If $\operatorname{spec}(JA)$ is simple, there exists a linear symplectic change of coordinates which brings the quadratic part of the Hamiltonian into normal form. Denoting the new coordinates by the same symbols as the old ones one has

$$\langle Aw, w \rangle = \sum_{i=1}^{n} \lambda_i (q_i^2 + p_i^2).$$

Definition 2.3.1. *A Hamiltonian H is in* Birkhoff normal form *up to order $m \geq 2$, if it is of the form*

$$H = N_2 + N_4 + \dots + N_m + H_{m+1} + \dots, \tag{2.3}$$

where the N_k, $2 \leq k \leq m$, *are homogeneous polynomials of order* k, *which are actually functions of* $q_1^2 + p_1^2, \ldots, q_n^2 + p_n^2$, *and where* $H_{m+1} + \ldots$ *stands for (arbitrary) terms of order strictly greater than* m. *If this holds for any* m, *the Hamiltonian is said to be in* Birkhoff normal form *and the coordinates* $(q_i, p_i)_{1 \leq i \leq n}$ *are referred to as* Birkhoff coordinates.

Note that if a Hamiltonian H admits a Birkhoff normal form of order m, the coefficients of the expansion (2.3) up to order m are uniquely determined, as long as the normalizing transformation is of the form $\mathrm{id} + \ldots$. However, the normalizing transformation is by no means unique.

There are well known theorems guaranteeing the existence of a Birkhoff normal form up to order m assuming that the frequencies $\lambda_1, \ldots, \lambda_n$ satisfy certain nonresonance conditions - see e.g. Theorem 4.3 in [42]. However, in the case investigated in this thesis, one-dimensional FPU chains, the nonresonance conditions for $m = 4$ are not satisfied. We will thus not use these general theorems but rather show by explicit computations that arbitrary odd periodic and Dirichlet FPU chains admit a Birkhoff normal form up to order four, while even periodic FPU chains only admit a *resonant* fourth order normal form except in the Toda case. Birkhoff's original work can be found in [8].

2.4 Perturbed integrable systems

Since many interesting physical systems can be viewed as perturbations of an integrable Hamiltonian system, one is interested in whether the foliation of invariant tori of the unperturbed system can still be found in the perturbed system. The theorems of KAM and Nekhoroshev theory give some answers to this question.

We consider a Hamiltonian in action-angle coordinates $(\theta, I) \in \mathbb{T}^n \times D$ (where D is a bounded domain in $\mathbb{R}^n$) of the form

$$H = H_0(I) + H_\varepsilon(\theta, I) \tag{2.4}$$

with an unperturbed integrable Hamiltonian H_0 and a perturbation H_ε, which for simplicity we assume to be of the form $H_\varepsilon(\theta, I) = \varepsilon H_1(\theta, I)$. The unperturbed system is *nondegenerate*, if the frequencies vary with the actions locally in a one-to-one manner, i.e. if the frequency map

$$I \mapsto \omega(I) = \frac{\partial H_0}{\partial I}(I)$$

is a local diffeomorphism everywhere in D. This is equivalent to requiring that for any $I \in D$, the frequency map $\omega(I)$ satisfies *Kolmogorov's condition*

$$\det \frac{\partial \omega}{\partial I}(I) = \det \frac{\partial^2 H_0}{\partial I^2}(I) \neq 0. \tag{2.5}$$

Whereas a dense set of tori is destroyed and a generic Hamiltonian system therefore is not integrable [52] (first results in this direction are due to Poincaré

[59]), it was Kolmogorov's discovery that the *majority* of tori survives a small perturbation, namely those whose frequencies ω are not only nonresonant but *strongly nonresonant* in the sense that they satisfy a *diophantine* or *small divisor condition* of the form

$$|\langle k, \omega\rangle| \geq \frac{\alpha}{|k|^\tau} \quad \text{for all } k \in \mathbb{Z}^n \setminus \{0\}. \tag{2.6}$$

We denote for fixed $\tau > 0$ by Δ_α the set of all $\omega \in \mathbb{R}^n$ satisfying (2.6) for some given $\alpha > 0$. It can be shown that almost every $\omega \in \mathbb{R}^n$ belongs to some Δ_α in the sense that if $\tau > n-1$ one has the Lebesgue measure estimate $\text{meas}(\Omega \setminus \Delta_\alpha) = O(\alpha)$. The parameter α in (2.6) can however not be chosen arbitrarily small for a given perturbation H_ε, it has to satisfy a condition of the form $\alpha >> \sqrt{\varepsilon}$. For the statement of the theorem, we will fix α and thus obtain an upper bound on the parameter ε measuring the size of the perturbation. Moreover, from a bounded domain $\Omega \subset \mathbb{R}^n$ we define the subsets $\Omega_\alpha \subset \Omega$ by

$$\Omega_\alpha := \{\omega \in \Omega | \omega \in \Delta_\alpha, \text{dist}\,(\omega, \partial\Omega) \geq \alpha\}.$$

It can be shown that these sets Ω_α are *Cantor sets*, i.e. that they are closed, perfect, and nowhere dense, and that they satisfy the same Lebesgue measure estimate as the sets Δ_α, if the boundary of Ω is piecewise smooth.

We now state the main theorem of Kolmogorov, Arnol'd, and Moser [43, 3, 53]. For simplicity, we cite its version from [42] for systems in action-angle coordinates, and not the version for systems given in rectangular coordinates, which we will apply in this thesis. The latter can be found in [62] or [74] - its main condition, namely the nondegeneracy of the frequency map of the unperturbed system, is the same as in the original version for systems in action-angle coordinates.

Theorem 2.4.1. *Suppose the Hamiltonian*

$$H = H_0 + H_\varepsilon$$

is real analytic on the closure of $\mathbb{T}^n \times D$, *where* D *is a bounded domain in* $\mathbb{R}^n$. *If the frequency map of the integrable Hamiltonian is a diffeomorphism* $D \to \Omega$, *then there exists a constant* $\delta > 0$ *such that for*

$$|\varepsilon| < \delta\alpha^2$$

all Kronecker tori $(\mathbb{T}^n, \omega)$ *of the unperturbed system with* $\omega \in \Omega_\alpha$ *persist as Lagrangian tori, being only slightly deformed. Moreoever, they depend in a Lipschitz continuous way on* ω *and fill the phase space* $\mathbb{T}^n \times D$ *up to a set of measure* $O(\alpha)$.

For a proof of Theorem 2.4.1, besides the original references mentioned above we refer to Pöschel's papers [62, 65].

Let us remark that several of the conditions listed in Theorem 2.4.1 can be weakened. On the one hand, neither the unperturbed Hamiltonian nor the perturbation have to be real analytic, it suffices that they are differentiable of class C_l for sufficiently large l (depending on the dimension n). On the other hand, the nondegeneracy condition can be replaced by other conditions on the frequency map $I \mapsto \omega(I)$, e.g. by isoenergetic nondegeneracy. To be precise, we call the unperturbed Hamiltonian H_0 *isoenergetically nondegenerate* in D if for any $I \in D$,

$$\det \begin{pmatrix} \partial^2 H_0/\partial I^2 & \partial H_0/\partial I \\ \partial H_0/\partial I & 0 \end{pmatrix}(I) = \det \begin{pmatrix} \partial \omega(I)/\partial I & \omega(I) \\ \omega(I) & 0 \end{pmatrix} \neq 0. \qquad (2.7)$$

For a KAM theorem for systems satisfying (2.7) we refer to [1]. Note that the conditions (2.5) and (2.7) are independent, i.e. none of the two conditions implies the other one (see [10] for an illustration of this fact by examples).

Further, it has been shown that instead of nondegeneracy or isoenergetic nondegeneracy, it suffices to require that the image of the frequency map $I \mapsto \omega(I)$ does not lie in any hyperplane in $\mathbb{R}^n$ passing through the origin (see e.g. [10] or [77]). Rüssmann then found criteria for this requirement involving terms of arbitrarily high order of the Taylor expansion of $H_0(I)$. In terms of the Taylor coefficients up to order 2, Rüssmanns nondegeneracy condition means that for any $I \in D$, the columns of the Hessian of H_0 are complementary in $\mathbb{R}^n$ to $\omega(I)$, i.e. the $n \times (n+1)$-matrix

$$\left(\partial \omega(I)/\partial I \middle| \omega(I) \right) \qquad (2.8)$$

has rank n. One easily sees that both (2.5) and (2.7) imply that the matrix (2.8) has rank n. There exist many further developments in KAM theory, in particular relaxations of other assumptions of the classical KAM theorem - for an extensive discussion see e.g. [48].

Even though Theorem 2.4.1 or its extensions discussed above guarantee the persistence of a majority of the invariant tori of an integrable system under a sufficiently small perturbation, the theorem is of somewhat probabilistic nature, since the tori whose frequencies do not satisfy the small divisor condition (2.6) are dense among all invariant tori of the unperturbed system (as are the tori who do satisfy (2.6)). It was Nekhoroshev [54, 55, 56] who first proved a type of result providing bounds on the variation of *all* orbits over a finite, but *exponentially long* time interval, under a slightly stronger assumption than the nondegeneracy of the KAM theorem, namely "*steepness*". In [56], he gives algebraic criteria for convexity, involving the coeffients of higher order terms of the Taylor expansion of $H_0(I)$, and in [37], Il'yaschenko gives an (analytic) criterion for steepness.

Because Nekhoroshev's notion of "steepness" seems rather difficult to check for a given system, we do not cite his original result here, but a more recent version by Pöschel [64] applicable to an elliptic equilibrium. Instead of Nekhoroshev's steepness, the assumption on the unperturbed Hamiltonian is convexity. Precisely, we assume that the Hamiltonian H is given in rectangular coordinates $(x_i, y_i)_{1 \le i \le n}$ by

$$H = \langle \alpha, I \rangle + \frac{1}{2} \langle QI, I \rangle + B(I) + P(x, y),$$

where as above $I = (I_1, \ldots, I_n)$ with $I_j = \frac{1}{2}(x_j^2 + y_j^2)$ for $1 \leq j \leq n$, the term B is of order 3 in I and absent for $4 \leq l \leq 5$ and $P = O_{l+1}(x, y)$ is of order $l + 1$ in x and y. The integrable unperturbed Hamiltonian

$$H_0 = \langle \alpha, I \rangle + \frac{1}{2}\langle QI, I \rangle + B(I)$$

is assumed to be convex in I for small I, which is equivalent to Q being positive definite. The following theorem was first proved in [17] or [57]; Pöschel gave a new proof based on a method by Lochak [49]. Let $|I| = |I_1| + \ldots + |I_n|$.

Theorem 2.4.2. *Suppose Q is positive definite. Then for every orbit of the Hamiltonian H with $|I(0)| < \delta^2$ sufficiently small one has*

$$|I(t) - I(0)| < c\delta^{2+\lambda a} \quad \textit{for} \quad |t| < \frac{1}{|\alpha|} e^{d\delta^{-\lambda a}} \tag{2.9}$$

with

$$a = \frac{1}{2n}, \quad \lambda = l - 3,$$

where the constants c and d only depend on Q and the dimension n.

It is likely that Theorem 2.4.2 is also true if the unperturbed Hamiltonian is only quasi-convex instead of convex; Pöschel proved such a theorem in [63], however only for systems in action-angle coordinates and not for systems in rectangular coordinates, as we typically have in a neighborhood of an isolated equilibrium.

Since this assumption of quasi-convexity instead of convexity considerably extends the applications of Theorem 2.4.2 to a given system, let us quantify the notion of quasi-convexity. Let $Q(I) := \frac{\partial^2 H_0}{\partial I^2}(I)$ be the Hessian of the integrable unperturbed Hamiltonian H_0 at $I \in D$, and let $l, m > 0$. We define H_0 to be *m-convex*, if the inequality

$$\langle Q(I)\xi, \xi \rangle \geq m\|\xi\|^2 \quad \forall \, \xi \in \mathbb{R}^n \tag{2.10}$$

holds at every point I in D. As a generalization, we define H_0 to be *l, m-quasi-convex*, if at every point $I \in D$ either (2.10) or

$$|\langle \omega(I), \xi \rangle| > l\|\xi\| \quad \forall \, \xi \in \mathbb{R}^n \tag{2.11}$$

holds. Note that although quasi-convexity is a generalization of convexity, it is still stronger than the original notion of steepness. We will however not investigate for the systems considered in this thesis for which parameter values the definition of quasi-convexity is fulfilled - this remains an open question.

Chapter 3

Normal form computations

In this chapter we perform all computations necessary for obtaining the various normal forms for FPU chains with the three different types of parities and boundary conditions, as stated in Theorems 1.2.1, 1.3.1, and 1.4.1. Whereas we obtain Birkhoff normal forms of order four for odd periodic and Dirichlet chains, we obtain a *resonant* normal form of order four for even periodic chains, which we then show to be completely integrable.

3.1 Odd periodic chains

We start with formula (1.5) for the Hamiltonian H_V of periodic FPU chains in the "physical" coordinates $(q_n, p_n)_{1 \leq n \leq N}$. We will prove Theorem 1.2.1 by four (explicit) transformations. The third and fourth of these transformations come up "naturally" if one tries to eliminate the non-normal form terms step by step, i.e. in increasing order. The first transformation - introducing relative coordinates - is an "immediate" consequence of neglecting the motion of the center of mass coordinate. The second transformation however seems unmotivated; it comes up by following our procedure of constructing Birkhoff coordinates for the periodic Toda lattice [33] which in turn is a finite-dimensional analogue of a method used by Kappeler and Pöschel [42] for the periodic KdV equation. We have put some of the computations which are necessary for these first two steps in Appendix A.

We first introduce relative coordinates,

$$v_i := q_{i+1} - q_i \ (1 \leq i \leq N-1) \quad \text{and} \quad v_N := \frac{1}{N} \sum_{i=1}^{N} q_i, \tag{3.1}$$

and denote by $(u_i)_{1 \leq i \leq N}$ the corresponding conjugate variables. It turns out that $u_N = N \cdot P = \sum_{i=1}^{N} p_i$ and $q_{N+1} - q_N = -\sum_{k=1}^{N-1} v_k$. The Hamiltonian H_V in (1.5), when expressed in these coordinates, takes the form $H_V = \frac{NP^2}{2} + \tilde{H}_V$ with $\tilde{H}_V = H_u + H_v$, where H_u and H_v only depend on $u = (u_i)_{1 \leq i \leq N-1}$ and

$v = (v_i)_{1\le i\le N-1}$, respectively, and are given by

$$\begin{aligned} H_u &= \frac{1}{2}\left(u_1^2 + \sum_{l=1}^{N-2}(u_{l+1}-u_l)^2 + u_{N-1}^2\right), \\ H_v &= \frac{1}{2}\left(\sum_{k=1}^{N-1} v_k^2 + \left(\sum_{k=1}^{N-1} v_k\right)^2\right) + \frac{\alpha}{3!}\left(\sum_{k=1}^{N-1} v_k^3 - \left(\sum_{k=1}^{N-1} v_k\right)^3\right) \\ &\quad + \frac{\beta}{4!}\left(\sum_{k=1}^{N-1} v_k^4 + \left(\sum_{k=1}^{N-1} v_k\right)^4\right) + O(v^5). \end{aligned}$$

More details of this transformation to relative coordinates can be found in section A.1.

To bring $\tilde{H}_V = H_u + H_v$ into Birkhoff normal form up to order two we introduce new coordinates $(\xi_k, \eta_k)_{1\le k\le N-1}$. It turns out to be convenient to use complex notation, i.e. for $1 \le k \le N-1$

$$\begin{cases} \zeta_k = \frac{1}{\sqrt{2}}(x_k - iy_k), \\ \zeta_{-k} = \overline{\zeta_k} = \frac{1}{\sqrt{2}}(x_k + iy_k). \end{cases} \tag{3.2}$$

where the minus sign in the definition of ζ_k is chosen so that $d\zeta_k \wedge d\zeta_{-k} = i d\xi_k \wedge d\eta_k$. The vector $\zeta = (\zeta_k)_{1\le |k|\le N-1}$ is an element in the space

$$\mathcal{Z} := \{z = (z_k)_{1\le|k|\le N-1} \in \mathbb{C}^{2N-2} : z_{-k} = \overline{z_k} \quad \forall\, 1 \le k \le N-1\}. \tag{3.3}$$

Further, for the rest of this thesis, we introduce the notations

$$s_k := \sin\frac{k\pi}{N}, \quad \lambda_k := \sqrt{|s_k|} = \sqrt{\left|\sin\frac{k\pi}{N}\right|} \qquad (0 \le |k| \le N-1). \tag{3.4}$$

The proposed linear transformation $\mathcal{Z} \to \mathbb{R}^{2N-2}$, $\zeta \mapsto (v, u)$ is then defined by

$$u_1(\zeta) = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \lambda_k \zeta_k, \tag{3.5}$$

$$u_{l+1}(\zeta) - u_l(\zeta) = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \lambda_k e^{2\pi i lk/N} \zeta_k \quad (1 \le l \le N-2), \tag{3.6}$$

$$-u_{N-1}(\zeta) = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \lambda_k e^{2\pi i (N-1)k/N} \zeta_k \tag{3.7}$$

and

$$v_l(\zeta) = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \lambda_k e^{2\pi i lk/N} e^{-i\pi k/N} \zeta_k \quad (1 \le l \le N-1). \tag{3.8}$$

Note that (3.7) is actually a consequence of (3.5) and (3.6). As explained in [34] it follows from the construction of the Birkhoff map of the periodic Toda lattice that this map is a canonical isomorphism. In order to keep this thesis self-contained, we have included a proof of this fact in section A.2.

When expressed in the new coordinates, H_u and H_v take the form

$$\begin{aligned} H_u(\zeta) &= \frac{1}{2} G_2(\zeta), \\ H_v(\zeta) &= \frac{1}{2} G_2(\zeta) + \alpha G_3(\zeta) + \beta G_4(\zeta) + O(\zeta^5), \end{aligned}$$

where

$$G_2 := 2 \sum_{k=1}^{N-1} s_k \zeta_k \zeta_{-k}, \tag{3.9}$$

$$G_3 := \frac{1}{6\sqrt{N}} \sum_{(k,k',k'') \in K_3} (-1)^{(k+k'+k'')/N} \lambda_k \lambda_{k'} \lambda_{k''} \zeta_k \zeta_{k'} \zeta_{k''}, \tag{3.10}$$

$$G_4 := \frac{1}{24N} \sum_{(k,k',k'',k''') \in K_4} (-1)^{(k+k'+k''+k''')/N} \lambda_k \lambda_{k'} \lambda_{k''} \lambda_{k'''} \zeta_k \zeta_{k'} \zeta_{k''} \zeta_{k'''}, \tag{3.11}$$

with

$$\begin{aligned} K_3 := \{(k,k',k'') \in \mathbb{Z}^3 :\ & 1 \le |k|, |k'|, |k''| \le N-1 \\ & \text{and } k + k' + k'' \equiv 0 \bmod N\} \end{aligned} \tag{3.12}$$

and

$$\begin{aligned} K_4 := \{(k,k',k'',k''') \in \mathbb{Z}^4 :\ & 1 \le |k|, |k'|, |k''|, |k'''| \le N-1 \\ & \text{and } k + k' + k'' + k''' \equiv 0 \bmod N\}. \end{aligned} \tag{3.13}$$

Note that G_2, G_3, and G_4 are *independent* of α and β. In particular they already came up in [34] when we computed the Birkhoff normal form of the periodic Toda lattice. For a detailed derivation of the formulas for G_2, G_3, and G_4 see section A.1. Summarizing the results of this section we have that

$$\tilde{H}_V(\zeta) = G_2(\zeta) + \alpha G_3(\zeta) + \beta G_4(\zeta) + O(\zeta^5)$$

is in Birkhoff normal form up to order two. As a consequence, $\zeta = 0$ is an elliptic fixed point of the Hamiltonian $\tilde{H}_V$.

We now begin by transforming $\tilde{H}_V(\zeta)$ into its Birkhoff normal form up to order four. Here we follow a standard procedure - see e.g. section 14 in [42]. The phase space $\mathcal{Z}$, defined in (3.3), is endowed with the Poisson bracket

$$\{F, G\} = i \sum_{1 \le |k| \le N-1} \sigma_k \frac{\partial F}{\partial \zeta_k} \frac{\partial G}{\partial \zeta_{-k}},$$

where $\sigma_k = \operatorname{sgn}(k)$ is the sign of k. The Hamiltonian vector field X_F associated to the Hamiltonian F is then given by $X_F = i \sum_{1\le|k|\le N-1} \sigma_k \frac{\partial F}{\partial \zeta_{-k}} \frac{\partial}{\partial \zeta_k}$. With a first canonical transformation we want to eliminate the third order term αG_3 in $\tilde{H}_V(\zeta)$. By a by now standard precedure we construct such a canonical transformation on the phase space $\mathcal{Z}$ as the time-1-map $\Psi_1 := X^t_{\alpha F_3}|_{t=1}$ of the flow $X^t_{\alpha F_3}$ of a real analytic Hamiltonian αF_3 which is a homogeneous polynomial in ζ_k $(1 \le |k| \le N-1)$ of degree 3 and solves the homological equation

$$\{G_2, \alpha F_3\} + \alpha G_3 = 0. \tag{3.14}$$

To simplify the notation we momentarily write F instead of αF_3 and H instead of $\tilde{H}_V$. Assuming for the moment that (3.14) can be solved and that X^t_F is defined for $0 \le t \le 1$ in some neighbourhood of the origin in $\mathcal{Z}$, we can use Taylor's formula to expand $H \circ X^t_F$ around $t = 0$,

$$\begin{aligned} H \circ X^t_F &= H \circ X^0_F + \int_0^t \frac{d}{ds}(H \circ X^s_F) ds \\ &= H + \int_0^t \{H, F\} \circ X^s_F \, ds \\ &= H + t\,\{H, F\} + \int_0^t ds \int_0^s ds' \frac{d}{ds'}(\{H, F\} \circ X^{s'}_F) \\ &= H + t\,\{H, F\} + \int_0^t (t-s)\{\{H, F\}, F\} \circ X^s_F \, ds. \end{aligned} \tag{3.15}$$

When evaluating this expression at $t = 1$, one gets

$$\begin{aligned} H \circ \Psi_1 = G_2 + \{G_2, F\} &+ \int_0^1 (1-t)\{\{G_2, F\}, F\} \circ X^t_F dt \\ &+\alpha G_3 + \int_0^1 \{\alpha G_3, F\} \circ X^t_F \, dt + \beta G_4 + O(\zeta^5). \end{aligned}$$

Using that $\{G_2, F\} + \alpha G_3 = 0$, the latter expression is simplified and we get

$$H \circ \Psi_1 = G_2 + \int_0^1 t\,\{\alpha G_3, F\} \circ X^t_F \, dt + \beta G_4 + O(\zeta^5).$$

Integrating by parts once more and taking into account that $F \equiv \alpha F_3$ is homogeneous of degree 3 one obtains, in view of (3.15),

$$\tilde{H}_V \circ \Psi_1 = G_2 + \frac{1}{2}\{\alpha G_3, \alpha F_3\} + \beta G_4 + O(\zeta^5). \tag{3.16}$$

Note that $\{G_3, F_3\}$ is homogeneous of order 4. Hence our first step is achieved. It remains to solve (3.14). Since G_3 contains only monomials with $(k, k', k'') \in K_3$ (cf (3.12)), also F_3 need only contain such monomials,

$$F_3 = \sum_{(k,k',k'')\in K_3} F^{(3)}_{kk'k''} \zeta_k \zeta_{k'} \zeta_{k''} \tag{3.17}$$

which leads to

$$\begin{aligned}\{G_2, F_3\} &= 2i \sum_{1\le|k|\le N-1} s_k \zeta_{-k} \frac{\partial F_3}{\partial \zeta_{-k}} \\ &= -2i \sum_{(k,k',k'')\in K_3} (s_k + s_{k'} + s_{k''}) F^{(3)}_{kk'k''} \zeta_k \zeta_{k'} \zeta_{k''}. \end{aligned} \tag{3.18}$$

The following result is due to Beukers and Rink (cf. [70, 73]):

Lemma 3.1.1. *For any* $(k, k', k'') \in K_3$*,*

$$s_k + s_{k'} + s_{k''} \neq 0.$$

Let us remark that Lemma 3.1.1 also follows from the integrability of the Toda lattice (cf. [34]). We include the self-contained proof due to Beukers and Rink.

Proof. Suppose that $(k, k', k'') \in K_3$ satisfies $s_k + s_{k'} + s_{k''} = 0$. It follows from $k + k' + k'' \equiv 0 \mod N$ that either $s_{k''} = -s_{k+k'}$ or $s_{k''} = s_{k+k'}$, according to whether $k + k' + k'' \equiv 0$ or $k + k' + k'' \equiv N \mod 2N$.

In the first case, it follows that

$$2i \sin \frac{k\pi}{N} + 2i \sin \frac{k'\pi}{N} - 2i \sin\left(\frac{k\pi}{N} + \frac{k'\pi}{N}\right) = 0. \tag{3.19}$$

Setting $x := e^{\frac{ik\pi}{N}}$ and $y := e^{\frac{ik'\pi}{N}}$, one can rewrite (3.19) as

$$0 = x - \frac{1}{x} + y - \frac{1}{y} - xy + \frac{1}{xy} = (1 - x)(1 - y)(1 - xy)\frac{1}{xy}. \tag{3.20}$$

It follows that any solution of (3.20) contradicts the assumption $(k, k', k'') \in K_3$, in particular $1 \le |k|, |k'|, |k''| \le N - 1$. Indeed, solutions with $x = 1$ (i.e. $k \equiv 0 \mod 2N$), $y = 1$ (i.e. $k' \equiv 0 \mod 2N$), or $xy = 1$ (i.e. $k + k' \equiv 0 \mod 2N$ and thus $k'' \equiv 0 \mod 2N$), contradict this assumption.

In the second case, we have instead of (3.19)

$$2i \sin \frac{k\pi}{N} + 2i \sin \frac{k'\pi}{N} + 2i \sin\left(\frac{k\pi}{N} + \frac{k'\pi}{N}\right) = 0. \tag{3.21}$$

With x, y as above, it now follows from (3.21) that

$$0 = x - \frac{1}{x} + y - \frac{1}{y} + xy - \frac{1}{xy} = -(1 + x)(1 + y)(1 - xy)\frac{1}{xy}.$$

Again we conclude that any solution of (3.21) contradicts the assumption $1 \le |k|, |k'|, |k''| \le N - 1$. Indeed, solutions with $x = -1$ (i.e. $k \equiv N \mod 2N$), $y = -1$ (i.e. $k' \equiv N \mod 2N$), or $xy = 1$ (i.e. $k + k' \equiv 0 \mod 2N$ and thus $k'' \equiv N \mod 2N$), contradict this assumption. $\square$

By Lemma 3.1.1, one can define F_3 as follows

$$iF^{(3)}_{kk'k''} := \begin{cases} \frac{G^{(3)}_{kk'k''}}{2(s_k+s_{k'}+s_{k''})} & (k,k',k'') \in K_3, \\ \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{G_2, \alpha F_3\} + \alpha G_3 = 0$. Written more explicitly, the nonzero coefficients of F_3 are

$$iF^{(3)}_{kk'k''} = \frac{(-1)^{(k+k'+k'')/N}}{6\sqrt{N}} \frac{\sqrt{|\sin\frac{k\pi}{N}\sin\frac{k'\pi}{N}\sin\frac{k''\pi}{N}|}}{2\sin\frac{k\pi}{N} + 2\sin\frac{k'\pi}{N} + 2\sin\frac{k''\pi}{N}}. \tag{3.22}$$

In a second step we normalize the 4th order term $\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}$ in (3.16). We decompose this sum into its contribution to the Birkhoff normal form and the rest, to be transformed away in a moment. Let us first compute $\{G_3, F_3\}$ in a more explicit form:

$$\begin{aligned}
\{G_3, F_3\} &= i \sum_{1\le|k|\le N-1} \sigma_k \frac{\partial G_3}{\partial \zeta_k}\frac{\partial F_3}{\partial \zeta_{-k}} = \sum_{1\le|k|\le N-1} \sigma_k \frac{\partial G_3}{\partial \zeta_k}\frac{\partial (iF_3)}{\partial \zeta_{-k}} \\
&= \frac{1}{N} \sum_{1\le|k|\le N-1} \sigma_k \left(\frac{3}{6} \sum_{\substack{1\le|l|,|m|\le N-1, \\ l+m=-k+rN}} (-1)^r \lambda_k\lambda_l\lambda_m \zeta_l\zeta_m \right) \\
&\qquad \cdot \left(\frac{3}{6} \sum_{\substack{1\le|l'|,|m'|\le N-1, \\ l'+m'=k+r'N}} (-1)^{r'} \frac{\lambda_k\lambda_{l'}\lambda_{m'}}{2(s_{-k}+s_{l'}+s_{m'})} \zeta_{l'}\zeta_{m'} \right) \\
&= \frac{1}{8N} \sum_{1\le|k|\le N-1} \sum_{\substack{1\le|l|,|m|,|l'|,|m'|\le N-1 \\ l+m-rN=-k \\ l'+m'-r'N=k}} (-1)^{r+r'} \frac{s_k\lambda_l\lambda_m\lambda_{l'}\lambda_{m'}}{s_{-k}+s_{l'}+s_{m'}} \zeta_l\zeta_m\zeta_{l'}\zeta_{m'},
\end{aligned}$$

where for the latter equality we used that $\sigma_k\lambda_k^2 = s_k$. Setting

$$\varepsilon_{lml'm'} := \frac{l+m+l'+m'}{N} \tag{3.23}$$

one then gets

$$\begin{aligned}
\{G_3, F_3\} &= \frac{1}{8N} \sum_{1\le|k|\le N-1} \sum_{\substack{l+m\equiv -k \bmod N \\ l'+m'\equiv k \bmod N}} (-1)^{\varepsilon_{lml'm'}} \frac{\lambda_l\lambda_m\lambda_{l'}\lambda_{m'}}{-1+(s_{l'}+s_{m'})/s_k} \zeta_l\zeta_m\zeta_{l'}\zeta_{m'} \\
&= \frac{1}{8N} \sum_{k=1}^{N-1} \sum_{\substack{l+m\equiv -k \bmod N \\ l'+m'\equiv k \bmod N}} (-1)^{\varepsilon_{lml'm'}} \frac{\lambda_l\lambda_m\lambda_{l'}\lambda_{m'}}{-1+(s_{l'}+s_{m'})/s_k} \zeta_l\zeta_m\zeta_{l'}\zeta_{m'}
\end{aligned}$$

$$+\frac{1}{8N}\sum_{k=1}^{N-1}\sum_{\substack{l+m\equiv k \bmod N\\ l'+m'\equiv -k \bmod N}}(-1)^{\varepsilon_{lml'm'}}\frac{\lambda_l\lambda_m\lambda_{l'}\lambda_{m'}}{-1-(s_{l'}+s_{m'})/s_k}\zeta_l\zeta_m\zeta_{l'}\zeta_{m'}$$

$$=\frac{1}{8N}\sum_{k=1}^{N-1}\sum_{\substack{l+m\equiv -k \bmod N\\ l'+m'\equiv k \bmod N}}\left(\frac{1}{-1+(s_{l'}+s_{m'})/s_k}+\frac{1}{-1-(s_l+s_m)/s_k}\right)$$
$$\cdot(-1)^{\varepsilon_{lml'm'}}\lambda_l\lambda_m\lambda_{l'}\lambda_{m'}\zeta_l\zeta_m\zeta_{l'}\zeta_{m'}.$$

Note that for $k = l' + m' + r'N$ with $1 \le k \le N-1$ and $r' \in \mathbb{Z}$ we have

$$s_k = |s_{l'+m'}|.$$

Introduce[1] for any $(l, m, l', m') \in K_4$

$$c_{lml'm'} = \begin{cases} \frac{1}{-1+\frac{s_{l'}+s_{m'}}{|s_{l'+m'}|}} - \frac{1}{1+\frac{s_l+s_m}{|s_{l+m}|}} & \text{if } l+m \not\equiv 0 \bmod N,\\ 0 & \text{otherwise.}\end{cases} \tag{3.24}$$

We then get

$$\frac{\alpha^2}{2}\{G_3, F_3\} = \frac{\alpha^2}{16N}\sum_{(l,m,l',m')\in K_4} c_{lml'm'}(-1)^{\varepsilon_{lml'm'}}\lambda_l\lambda_m\lambda_{l'}\lambda_{m'}\zeta_l\zeta_m\zeta_{l'}\zeta_{m'}. \tag{3.25}$$

Combined with formula (3.11) for G_4, the quantity $\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}$ becomes

$$\frac{1}{24N}\sum_{(k,k',k'',k''')\in K_4}(-1)^{\varepsilon_{kk'k''k'''}}(\beta+\frac{3\alpha^2}{2}c_{kk'k''k'''})\lambda_k\lambda_{k'}\lambda_{k''}\lambda_{k'''}\zeta_k\zeta_{k'}\zeta_{k''}\zeta_{k'''}. \tag{3.26}$$

We now decompose (3.26) into its contribution to the Birkhoff normal form of H_V and the rest, and we denote by π_N the projection onto the former one, whereas the latter one will be (partially) transformed away by a second transformation Ψ_2.

Lemma 3.1.2. *The normal form part of $\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}$ is given by*

$$\pi_N\Big(\ \beta\ G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}\Big)$$
$$=\frac{1}{4N}\left(\sum_{l=1}^{N-1}(\alpha^2+(\beta-\alpha^2)s_l^2)|\zeta_l|^4+2\sum_{1\le l\ne m\le N-1}(\beta-\alpha^2)s_l\ s_m|\zeta_l|^2|\zeta_m|^2\right). \tag{3.27}$$

Proof. The indices k, k', k'', k''' of the terms in $\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}$ contributing to the normal form satisfy $(k, k', k'', k''') \in K_4^N$, where

$$K_4^N := \{(k, k', k'', k''') \in K_4 |\ \exists\, 1 \le l \le m \le N-1 \text{ such that}$$
$$\{k, k', k'', k'''\} = \{l, -l, m, -m\}\}. \tag{3.28}$$

[1] To keep the formula for $c_{lml'm'}$ as simple as possible we have not symmetrized the coefficients $c_{lml'm'}$.

In the case $l = m$, $\{l, -l, l, -l\}$ in (3.28) is viewed as a set-like object whose two elements l and $-l$ each have multiplicity two.

We investigate $\pi_N(\beta G_4)$ and $\pi_N(\frac{\alpha^2}{2}\{G_3, F_3\})$ separately. Let us start with βG_4. We distinguish the cases $l = m$ and $l \neq m$ in K_4^N. For $l = m$, there are $\binom{4}{2} = 6$ distinct permutations of (k, k', k'', k''') in K_4^N, whereas for $l \neq m$, all $4! = 24$ permutations of $(l, m, -l, -m)$ are distinct. Hence we have

$$\begin{aligned}\pi_N(\beta G_4) &= \frac{\beta}{24N}\left(6\sum_{l=1}^{N-1} s_l^2|\zeta_l|^4 + 24\sum_{1\le l<m\le N-1} s_l\, s_m|\zeta_l|^2|\zeta_m|^2\right)\\ &= \frac{\beta}{4N}\left(\sum_{l=1}^{N-1} s_l^2|\zeta_l|^4 + 2\sum_{1\le l\neq m\le N-1} s_l\, s_m|\zeta_l|^2|\zeta_m|^2\right). \qquad (3.29)\end{aligned}$$

Now let us compute $\pi_N(\frac{\alpha^2}{2}\{G_3, F_3\})$. We have to single out the matches of (3.28) for which in addition the coefficient $c_{kk'k''k'''}$ in (3.25) does not vanish, i.e.

$$k + k' \not\equiv 0 \bmod N \text{ and } k + k' + k'' + k''' \equiv 0 \bmod N.$$

There are two quadruples (k, k', k'', k''') in K_4^N which satisfy these additional conditions,

$$\begin{matrix} k + k'' = 0 \\ k' + k''' = 0 \end{matrix} \qquad \text{or} \qquad \begin{matrix} k + k''' = 0 \\ k' + k'' = 0 \end{matrix}\;. \qquad (3.30)$$

In both cases, we have $s_{k''} + s_{k'''} = -(s_k + s_{k'})$, and therefore (3.24) reduces to

$$c_{kk'k''k'''} = \frac{-2|s_{k+k'}|}{|s_{k+k'}| + s_k + s_{k'}}. \qquad (3.31)$$

Note that (3.31) remains valid for $k + k' = N$, since in this case $s_{k+k'} = 0$ and $s_k + s_{k'} > 0$ as k and k' must satisfy $1 \le k, k' \le N - 1$, but not for $k + k' = 0$, since in this case $|s_{k+k'}| + s_k + s_{k'} = 0$.

We first compute the diagonal part of $\pi_N\left(\frac{1}{2}\{G_3, F_3\}\right)$. In this case, the two possibilities in (3.30) coincide and the solutions are

$$(k, k', k'', k''') = \begin{cases} (l, \;\; l, \; -l, -l) \\ (-l, -l, \;\; l, \;\; l) \end{cases}, \qquad (3.32)$$

where $1 \le l \le N - 1$. The sum of the coefficients $c_{kk'k''k'''}$ for the two cases listed in (3.32) is

$$c_{l,l,-l,-l} + c_{-l,-l,l,l} = -2|s_{2l}|\left(\frac{1}{|s_{2l}|+2s_l} + \frac{1}{|s_{2l}|-2s_l}\right) = \frac{-4s_{2l}^2}{s_{2l}^2-4s_l^2} = 4\cot^2\frac{l\pi}{N}.$$

We now turn to the off-diagonal part of $\pi_N\left(\frac{1}{2}\{G_3, F_3\}\right)$. The quadruples $(k, k', k'', k''') \in K_4$ satisfying (3.30) for given $\{l, m\} \subseteq \{1, \ldots, N - 1\}$ with

$l < m$, $(k, k') = (\pm l, \pm m)$, and $(k'', k''') = (\pm l, \pm m)$, are

$$(k, k', k'', k''') = \begin{cases} (l, & m, & -l, & -m) \\ (l, & -m, & -l, & m) \\ (-l, & m, & l, & -m) \\ (-l, & -m, & l, & m) \end{cases}. \tag{3.33}$$

The remaining matches are obtained from (3.33) by permuting the first and second or the third and fourth columns on the right hand side of (3.33), bringing the total number of all matches to $16 = 4 \cdot 4$. Note that by formula (3.31), these permutations leave the value of the coefficients $c_{kk'k''k'''}$ invariant. Taking the sum of the coefficients $c_{kk'k''k'''}$ for all the quadruples listed in (3.33), we obtain

$$\begin{aligned}
&4(c_{l,m,-l,-m} + c_{l,-m,-l,m} + c_{-l,m,l,-m} + c_{-l,-m,l,m}) \\
&\quad = -8\left(\frac{|s_{l+m}|}{|s_{l+m}| + s_l + s_m} + \frac{|s_{l-m}|}{|s_{l-m}| + s_l - s_m}\right. \\
&\qquad \left. + \frac{|s_{l-m}|}{|s_{l-m}| - s_l + s_m} + \frac{|s_{l+m}|}{|s_{l+m}| - s_l - s_m}\right) \\
&\quad = -16\left(\frac{s_{l-m}^2}{s_{l-m}^2 - (s_l - s_m)^2} + \frac{s_{l+m}^2}{s_{l+m}^2 - (s_l + s_m)^2}\right) \\
&\quad = \frac{-16(2s_{l-m}^2 s_{l+m}^2 - s_{l-m}^2 (s_l + s_m)^2 - s_{l+m}^2 (s_l - s_m)^2)}{s_{l-m}^2 s_{l+m}^2 + (s_l - s_m)^2 (s_l + s_m)^2 - s_{l-m}^2 (s_l + s_m)^2 - s_{l+m}^2 (s_l - s_m)^2} \\
&\quad = -16,
\end{aligned}$$

since $s_{l-m}^2 s_{l+m}^2 = (s_l - s_m)^2 (s_l + s_m)^2$. Collecting terms, we thus have

$$\begin{aligned}
\pi_N\left(\frac{\alpha^2}{2}\{G_3, F_3\}\right) &= \frac{\alpha^2}{16N}\left(\sum_{l=1}^{N-1} 4\cos^2\frac{\pi l}{N}|\zeta_l|^4 - 16 \sum_{1 \le l < m \le N-1} s_l\, s_m |\zeta_l|^2 |\zeta_m|^2\right) \\
&= \frac{\alpha^2}{4N}\left(\sum_{l=1}^{N-1} (1 - s_l^2)|\zeta_l|^4 - 2 \sum_{1 \le l \ne m \le N-1} s_l\, s_m |\zeta_l|^2 |\zeta_m|^2\right).
\end{aligned} \tag{3.34}$$

Adding up (3.29) and (3.34), we obtain (3.27). $\square$

Now we want to remove [as much as possible of] the term $(\mathrm{Id} - \pi_N)(\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\})$ from the Hamiltonian (3.16), $\tilde{H}_V \circ \Psi_1$, by a second coordinate transformation Ψ_2. In view of formulas (3.11) and (3.25) for G_4 and $\frac{1}{2}\{G_3, F_3\}$, respectively, and in complete analogy to the first step we look for a transformation Ψ_2 of the form $\Psi_2 = X_{F_4}^t|_{t=1}$ with

$$F_4 = \sum_{(k,k',k'',k''') \in K_4 \setminus K_4^N} F_{kk'k''k'''}^{(4)} \zeta_k \zeta_{k'} \zeta_{k''} \zeta_{k'''}, \tag{3.35}$$

where $F^{(4)}_{\sigma(k,k',k'',k''')} = F^{(4)}_{(k,k',k'',k''')}$ for any permutation $\sigma(k,k',k'',k''')$ of the quadruple $(k,k',k'',k''') \in K_4 \setminus K_4^N$. We would like to determine the coefficients of F_4 in such a way that

$$\{G_2, F_4\} = -(\mathrm{Id} - \pi_N)\left(\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}\right). \tag{3.36}$$

As in (3.18) one gets

$$\{G_2, F_4\} = -i \sum_{(k,k',k'',k''') \in K_4 \setminus K_4^N} (s_k + s_{k'} + s_{k''} + s_{k'''}) F^{(4)}_{kk'k''k'''} \zeta_k \zeta_{k'} \zeta_{k''} \zeta_{k'''}, \tag{3.37}$$

and equation (3.36) combined with (3.26) leads to

$$\begin{aligned} &i(s_k + s_{k'} + s_{k''} + s_{k'''}) F^{(4)}_{kk'k''k'''} \\ &\quad = \frac{1}{24N}(-1)^{\varepsilon_{kk'k''k'''}}\left(\beta + \frac{3\alpha^2}{2} c^S_{kk'k''k'''}\right) \cdot \lambda_k \lambda_{k'} \lambda_{k''} \lambda_{k'''} \end{aligned} \tag{3.38}$$

for any quadruple (k,k',k'',k''') in $K_4 \setminus K_4^N$. Here $c^S_{kk'k''k'''}$ denotes the symmetrized version of $c_{kk'k''k'''}$,

$$c^S_{kk'k''k'''} := \frac{1}{4!} \sum_{\sigma \in S_4} c_{\sigma(k,k',k'',k''')}. \tag{3.39}$$

The following lemma due to Beukers and Rink (cf. [70]) determines the quadruples $(k,k',k'',k''') \in K_4 \setminus K_4^N$ for which $s_k + s_{k'} + s_{k''} + s_{k'''} = 0$. Let us introduce

$$K_4^{res} := K^+_{res} \cup K^-_{res} \subseteq K_4$$

where

$$\begin{aligned} K^{\pm}_{res} := \Big\{ &(k,k',k'',k''') \in K_4 \,|\, \exists\, l \in \mathbb{N} : 1 \le l \le \frac{N}{4} \text{ so that} \\ &\{k,k',k'',k'''\} = \{\pm l, \pm l \mp N, \frac{N}{2} \mp l, -\frac{N}{2} \mp l\} \Big\}. \end{aligned}$$

Note that if N is odd, then $K_4^{res} = \emptyset$.

Lemma 3.1.3. *Let $(k_1,k_2,k_3,k_4) \in K_4 \setminus K_4^N$. Then*

$$s_k + s_{k'} + s_{k''} + s_{k'''} = 0 \textit{ if and only if } (k,k',k'',k''') \in K_4^{res}.$$

In particular, if N is odd, then $s_k + s_{k'} + s_{k''} + s_{k'''} \neq 0$.

For the convenience of the reader a detailed proof of Lemma 3.1.3 is given in Appendix B. It is likely that Lemma 3.1.3 also can be proved using the integrability of the Toda lattice (cf. [34]), which would be a remarkable fact insofar as it would be a proof of a number-theoretic fact with methods of the theory of dynamical systems.

By Lemma 3.1.3, if N is odd, (3.38) can be solved for any $(k,k',k'',k''') \in K_4 \setminus K_4^N$ determining the coefficients $F^{(4)}_{kk'k''k'''}$ with $(k,k',k'',k''') \in K_4 \setminus K_4^N$ in such a way that $F^{(4)}_{\sigma(k,k',k'',k''')} = F^{(4)}_{(k,k',k'',k''')}$ for any permutation $\sigma(k,k',k'',k''')$ of $(k,k',k'',k''') \in K_4 \setminus K_4^N$. With this choice of F_4 the canonical transformation Ψ_2 is then defined by $X^t_{F_4}|_{t=1}$. Composing Ψ_1 and Ψ_2, we obtain the transformation $\Xi := \Psi_1 \circ \Psi_2$. We have proved the following

Proposition 3.1.4. *Assume that $N \geq 3$ is odd. The real analytic symplectic coordinate transformation $\zeta = \Xi(z)$, defined in a neighborhood of the origin in $\mathcal{Z}$, transforms the Hamiltonian $\tilde{H}_V$ into its Birkhoff normal form up to order 4. More precisely,*

$$\tilde{H}_V \circ \Xi = G_2 + \pi_N \left(\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\} \right) + O(z^5), \tag{3.40}$$

with G_2 and $\pi_N(\beta G_4 + \frac{\alpha^2}{2}\{G_3,F_3\})$ given by (3.9) and (3.27), respectively.

Theorem 1.2.1 can now be proved easily.

Proof of Theorem 1.2.1. Proposition 3.1.4 provides the Taylor series expansion of $\tilde{H}_V$ in terms of the actions $I = (I_k)_{1\leq k\leq N-1}$ given by (1.6). More precisely, $\tilde{H}_V \circ \Xi = H_{\alpha,\beta}(I) + O(z^5)$, where $H_{\alpha,\beta}(I)$ is defined by

$$2\sum_{k=1}^{N-1} s_k I_k + \frac{1}{4N}\sum_{k=1}^{N-1}(\alpha^2 + (\beta-\alpha^2)s_k^2)I_k^2 + \frac{\beta-\alpha^2}{2N}\sum_{\substack{l\neq m \\ 1\leq l,m\leq N-1}} s_l s_m I_l I_m. \tag{3.41}$$

This proves Theorem 1.2.1. □

3.2 Even periodic chains

Now we assume that N is even. To obtain the normal form of the FPU Hamiltonian as claimed in Theorem 1.3.1 we continue the investigations of the previous section. According to Lemma 3.1.3, equation (3.38) might have no solution $F^{(4)}_{kk'k''k'''}$ for $(k,k',k'',k''') \in K_4^{res}$. We first compute the projection $\pi_{res}(\beta G_4 + \frac{\alpha^2}{2}\{G_3,F_3\})$ of $\beta G_4 + \frac{\alpha^2}{2}\{G_3,F_3\}$ onto those terms which are indexed by quadruples $(k,k',k'',k''') \in K_4^{res}$, i.e. the projection onto the resonant non-normal form part of $\beta G_4 + \frac{\alpha^2}{2}\{G_3,F_3\}$.

Lemma 3.2.1. *Assume that N is even. The resonant non-normal form part of $\beta G_4 + \frac{\alpha^2}{2}\{G_3,F_3\}$ is given by*

$$\pi_{res}\left(\beta G_4 + \frac{\alpha^2}{2}\{G_3,F_3\}\right) = -\frac{\beta-\alpha^2}{4N}(R + R_{\frac{N}{4}}) \tag{3.42}$$

where

$$R = 2\sum_{1\leq l<\frac{N}{4}} s_{2l}\left(\zeta_l\zeta_{-N+l}\zeta_{\frac{N}{2}-l}\zeta_{-\frac{N}{2}-l} + \zeta_{-l}\zeta_{N-l}\zeta_{\frac{N}{2}+l}\zeta_{-\frac{N}{2}+l}\right) \tag{3.43}$$

and

$$R_{\frac{N}{4}} = \begin{cases} \frac{1}{2}\left(\zeta^2_{\frac{N}{4}}\zeta^2_{-\frac{3N}{4}} + \zeta^2_{\frac{3N}{4}}\zeta^2_{-\frac{N}{4}}\right) & \textit{if } \frac{N}{4} \in \mathbb{N} \\ 0 & \textit{otherwise.} \end{cases} \tag{3.44}$$

Proof. Consider the formula (3.26) for $\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}$. At this point we need to consider the symmetrized version (3.39) of the coefficients $c_{klk'l'}$ defined by (3.24). We claim that for any $(k_1, k_2, k_3, k_4) \in K_4^{res}$

$$c^S_{k_1k_2k_3k_4} = \frac{1}{4!}\sum_{\sigma\in S_4} c_{k_{\sigma(1)}k_{\sigma(2)}k_{\sigma(3)}k_{\sigma(4)}} = -\frac{2}{3}. \tag{3.45}$$

Observe that $c_{k_1k_2k_3k_4}$ is invariant under the transpositions $k_1 \leftrightarrow k_2$ and $k_3 \leftrightarrow k_4$. Hence (3.45) follows once we prove that

$$4(c_{k_1k_2k_3k_4} + c_{k_1k_3k_2k_4} + c_{k_1k_4k_2k_3} + c_{k_2k_4k_1k_3} + c_{k_2k_3k_1k_4} + c_{k_3k_4k_1k_2}) = -16. \tag{3.46}$$

Note that any element $(k_1, k_2, k_3, k_4) \in K_4^{res}$ is, mod $2N$, a permutation of an element of the form $(l, -N+l, N/2-l, -N/2-l)$ with $1 \le |l| \le N/4$. For such quadruples one gets by a straightforward computation

$$c_{k_1k_2k_3k_4} + c_{k_3k_4k_1k_2} = -2 - 2 = -4$$

and, with $c_l = \cos\frac{l\pi}{N}$,

$$c_{k_1k_3k_2k_4} + c_{k_2k_4k_1k_3} = -\frac{4}{2+2(s_l+c_l)} - \frac{4}{2-2(s_l+c_l)} = -\frac{4}{s_{2l}}$$

as well as

$$c_{k_1k_4k_2k_3} + c_{k_2k_3k_1k_4} = -\frac{4}{2+2(s_l-c_l)} - \frac{4}{2-2(s_l-c_l)} = \frac{4}{s_{2l}}.$$

Substituting these three identities into the left hand side of (3.46) leads to the claimed identity (3.46).

Moreover, by the definition (3.23) of $\varepsilon_{lml'm'}$ one has for any $(k_1, k_2, k_3, k_4) \in K_4^{res}$ and any $\sigma \in S_4$, that $\varepsilon_{\sigma(k_1,k_2,k_3,k_4)} = \pm 1$ and hence

$$(-1)^{\varepsilon_{\sigma(k_1,k_2,k_3,k_4)}} = -1.$$

Further,

$$\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4} = \left|\sin\frac{l\pi}{N}\cos\frac{l\pi}{N}\right| = \frac{1}{2}\left|\sin\frac{2l\pi}{N}\right| = \frac{1}{2}|s_{2l}|.$$

Combining all these computations we get

$$\pi_{res}\Big(\beta G_4 + \frac{\alpha^2}{2}\{G_3, F_3\}\Big)$$
$$= \frac{1}{24N}\sum_{(k_1,k_2,k_3,k_4)\in K_4^{res}} (-1)(\beta + \frac{3\alpha^2}{2}c^S_{k_1k_2k_3k_4})\lambda_{k_1}\lambda_{k_2}\lambda_{k_3}\lambda_{k_4}\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\zeta_{k_4}$$

$$
\begin{aligned}
&= -\frac{4!\,(\beta-\alpha^2)}{24N}\sum_{1\le l<\frac{N}{4}}\frac{s_{2l}}{2}\left(\zeta_l\zeta_{-N+l}\zeta_{\frac{N}{2}-l}\zeta_{-\frac{N}{2}-l}+\zeta_{-l}\zeta_{N-l}\zeta_{\frac{N}{2}+l}\zeta_{-\frac{N}{2}+l}\right)\\
&\quad \underbrace{-\frac{3!\,(\beta-\alpha^2)}{24N}\cdot\frac{2}{4}\left(\zeta^2_{\frac{N}{4}}\zeta^2_{-\frac{3N}{4}}+\zeta^2_{-\frac{N}{4}}\zeta^2_{\frac{3N}{4}}\right)}_{\text{only if }\frac{N}{4}\in\mathbb{N}} \qquad (3.47)\\
&= -\frac{\beta-\alpha^2}{4N}(R+R_{\frac{N}{4}}),
\end{aligned}
$$

with R and $R_{\frac{N}{4}}$ as defined by (3.43) and (3.44), respectively. Hence Lemma 3.2.1 is proved. □

By Lemma 3.2.1, if N is even, equation (3.38) can be solved for any quadruple $(k,k',k'',k''')\in K_4\setminus(K_4^N\cup K_4^{res})$ in such a way that $F^{(4)}_{\sigma(k,k',k'',k''')}=F^{(4)}_{(k,k',k'',k''')}$ for any permutation $\sigma(k,k',k'',k''')$ of (k,k',k'',k'''). With this choice of F_4 the canonical transformation Ψ_2 is then defined by $X^t_{F_4}|_{t=1}$. Composing Ψ_1 and Ψ_2, we obtain the transformation $\Xi:=\Psi_1\circ\Psi_2$ and have proved the following

Proposition 3.2.2. *Assume that N is even. The real analytic symplectic coordinate transformation $\zeta=\Xi(z)$, defined locally in a neighborhood of the origin $z=0$ in $\mathcal{Z}$, transforms the Hamiltonian $\tilde{H}_V$ into the resonant Birkhoff normal form up to order 4,*

$$\tilde{H}_V\circ\Xi = G_2+\pi_N\left(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\}\right)+\pi_{res}\left(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\}\right)+O(z^5),$$

with G_2, $\pi_N(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\})$, and $\pi_{res}(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\})$ given by (3.9), (3.27), and (3.42), respectively.

Proof of Theorem 1.3.1. We start with the formula for $\tilde{H}_V\circ\Xi$ given by Proposition 3.2.2 and treat the normal form terms $G_2+\pi_N(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\})$ and the resonant normal form terms $\pi_{res}(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\})$ separately. With the action variables $I=(I_k)_{1\le k\le N-1}$ defined by (1.6) we see that $G_2+\pi_N(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\})=H_{\alpha,\beta}(I)$, where $H_{\alpha,\beta}(I)$ is defined by (3.41). Concerning the term $\pi_{res}(\beta G_4+\frac{\alpha^2}{2}\{G_3,F_3\})$, we first express it in terms of the real variables $(x_k,y_k)_{1\le k\le N-1}$, related to the ζ_k's by $x_k=(\zeta_k+\zeta_{-k})/2$ and $y_k=(\zeta_{-k}-\zeta_k)/2i$. Note that

$$
\begin{aligned}
&\zeta_l\zeta_{-N+l}\zeta_{\frac{N}{2}-l}\zeta_{-\frac{N}{2}-l}+\zeta_{-l}\zeta_{N-l}\zeta_{-\frac{N}{2}+l}\zeta_{\frac{N}{2}+l}\\
&\quad= 2\,\mathrm{Re}\,(\zeta_l\zeta_{-N+l}\zeta_{\frac{N}{2}-l}\zeta_{-\frac{N}{2}-l})\\
&\quad= \frac{1}{2}\big((x_lx_{N-l}+y_ly_{N-l})(x_{\frac{N}{2}-l}x_{\frac{N}{2}+l}+y_{\frac{N}{2}-l}y_{\frac{N}{2}+l})\\
&\qquad\quad -(x_ly_{N-l}-x_{N-l}y_l)(x_{\frac{N}{2}-l}y_{\frac{N}{2}+l}-x_{\frac{N}{2}+l}y_{\frac{N}{2}-l})\big)\\
&\quad= 2\left(J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}\right), \qquad (3.48)
\end{aligned}
$$

where for any $1 \leq k \leq N-1$, J_k and M_k are given by (1.9). Hence R, given by (3.43), can be expressed in terms of J_k and M_k as follows

$$\begin{aligned} R(J,M) &= 2 \sum_{1 \leq l < \frac{N}{4}} s_{2l} \left(\zeta_l \zeta_{-N+l} \zeta_{\frac{N}{2}-l} \zeta_{-\frac{N}{2}-l} + \zeta_{-l} \zeta_{N-l} \zeta_{\frac{N}{2}+l} \zeta_{-\frac{N}{2}+l} \right) \\ &= 4 \sum_{1 \leq l < \frac{N}{4}} \sin \frac{2l\pi}{N} \left(J_l J_{\frac{N}{2}-l} - M_l M_{\frac{N}{2}-l} \right). \end{aligned} \tag{3.49}$$

Similarly, if $\frac{N}{4} \in \mathbb{N}$, one concludes from (3.48) that $R_{\frac{N}{4}}$, given by (3.44), can be expressed as

$$R_{\frac{N}{4}}(J,M) = \frac{1}{2} \left(\zeta_{\frac{N}{4}}^2 \zeta_{-\frac{3N}{4}}^2 + \zeta_{\frac{3N}{4}}^2 \zeta_{-\frac{N}{4}}^2 \right) = J_{\frac{N}{4}}^2 - M_{\frac{N}{4}}^2. \tag{3.50}$$

Theorem 1.3.1 now follows from the formulas (3.41), (3.49), and (3.50). □

We now turn to the proof of Theorem 1.3.2, i.e. the integrability of the truncated Hamiltonian (1.14). Denote by $\{\cdot,\cdot\}$ the standard Poisson bracket on $\mathbb{R}^{2N-2}$. In a straightforward way one computes the Poisson brackets between the variables $I, M, J, L \in \mathbb{R}^{N-1}$, given by (1.6) and (1.9) (cf. [15], p. 28):

Lemma 3.2.3. *The Poisson brackets between the variables I_k, J_k, M_k ($1 \leq k \leq N-1$) are given by*

$$\begin{aligned} \{I_l, I_k\} &= \{J_l, J_k\} = \{M_l, M_k\} = 0, & (3.51) \\ \{J_l, I_k\} &= -M_l(\delta_{kl} - \delta_{k+l,N}), & (3.52) \\ \{M_l, I_k\} &= J_l(\delta_{kl} - \delta_{k+l,N}), & (3.53) \end{aligned}$$

As a consequence, one obtains the following relations between the variables M_k, J_k, and L_k, $1 \leq k \leq N-1$:

$$\begin{aligned} \{M_k, J_l\} &= L_l(\delta_{k+l,N} - \delta_{kl}), \\ \{J_k, L_l\} &= M_k(\delta_{k+l,N} - \delta_{kl}), \\ \{L_k, M_l\} &= J_l(\delta_{k+l,N} - \delta_{kl}). \end{aligned}$$

First note that the list of functions of Theorem 1.3.2,

$$(I_k + I_{N-k})_{1 \leq k \leq \frac{N}{2}}, \ (I_k + I_{\frac{N}{2}+k})_{1 \leq k < \frac{N}{4}}, \ (K_k)_{1 \leq k \leq \frac{N}{4}}, \tag{3.54}$$

contains $N-1$ terms regardless whether $\frac{N}{4}$ is an integer or not. In addition, for any $1 \leq k < \frac{N}{4}$, the terms $I_k + I_{N-k}, I_{N/2-k} + I_{N/2+k}, I_k + I_{N/2+k}, K_k$ are functions of the eight variables $x_k, y_k, x_{N/2-k}, y_{N/2-k}, x_{N/2+k}, y_{N/2+k}, x_{N-k}$, and y_{N-k}, the term $I_{N/2}$ is a function of the two variables $x_{N/2}, y_{N/2}$, and, in the case $N/4 \in \mathbb{N}$, the terms $I_{\frac{N}{4}} + I_{\frac{3N}{4}}, K_{\frac{N}{4}}$ are functions of the four variables $x_{N/4}, y_{N/4}, x_{3N/4}, y_{3N/4}$. Hence we obtain a partition of the $2N-2$ variables $x_1, y_1, \ldots, x_{N-1}, y_{N-1}$ into $\llcorner \frac{N}{4} \lrcorner + 1$ pairwise disjoint sets of variables, and all Poisson brackets between variables of different sets of this partition vanish.

Lemma 3.2.4. *The $N-1$ functions listed in (3.54) are pairwise in involution.*

Proof. The functions in (3.54) depend on only one of the $\llcorner\frac{N}{4}\lrcorner+1$ pairwise disjoint sets of variables. As the Poisson brackets between terms depending on variables of different sets vanish, it remains to check that functions of (3.54) with the same k are in involution with each other. In view of the formulas (1.17) and (1.19) for K_l and $K_{N/4}$, respectively (recall that $J_j^2+M_j^2=I_jI_{N-j}$ for any $1\le j<\frac{N}{2}$), and taking into account that $(I_k)_{1\le k\le N-1}$ are pairwise in involution, this amounts to proving that for any $1\le l<\frac{N}{4}$,

$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_l+I_{N-l}\}=0, \tag{3.55}$$
$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_{\frac{N}{2}-l}+I_{\frac{N}{2}+l}\}=0, \tag{3.56}$$
$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_l+I_{\frac{N}{2}+l}\}=0, \tag{3.57}$$

and

$$\{J_{\frac{N}{4}}^2-M_{\frac{N}{4}}^2, I_{\frac{N}{4}}+I_{\frac{3N}{4}}\}=0. \tag{3.58}$$

First we note that by (3.52) and (3.53) one has for any $1\le l<\frac{N}{2}$

$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_l\}=-J_{\frac{N}{2}-l}M_l-M_{\frac{N}{2}-l}J_l \tag{3.59}$$

and

$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_{N-l}\}=J_{\frac{N}{2}-l}M_l+M_{\frac{N}{2}-l}J_l. \tag{3.60}$$

Since the right hand sides of (3.59) and (3.60) are invariant under exchanging l and $\frac{N}{2}-l$, the same must hold for the left hand sides, and we conclude that

$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_{\frac{N}{2}-l}\}=-J_{\frac{N}{2}-l}M_l-M_{\frac{N}{2}-l}J_l \tag{3.61}$$

and

$$\{J_lJ_{\frac{N}{2}-l}-M_lM_{\frac{N}{2}-l}, I_{\frac{N}{2}+l}\}=J_{\frac{N}{2}-l}M_l+M_{\frac{N}{2}-l}J_l. \tag{3.62}$$

The identities (3.55)-(3.57) now follow from the appropriate combinations of (3.59)-(3.62). In the same fashion, one concludes that (3.58) holds. □

Proof of Theorem 1.3.2. In view of Lemma 3.2.4, it remains to check that the quantities listed in (3.54) are functionally independent integrals. The independence is easy to verify, and the fact that they are conserved quantities follows from the formula (1.16), showing that H_V^{trunc} can be written as a function of them. □

3.3 Dirichlet chains

In this section we consider a chain with N' ($N'\ge 3$, not necessarily even) moving particles and fixed endpoints, i.e. with boundary conditions (1.3).

It has been observed that such a chain can be treated as an invariant subsytem of a periodic lattice with $N=2N'+2$ particles - see [73]: Let $T^*\mathbb{R}^N$ be endowed with the canonical symplectic structure and consider the linear map $S : T^*\mathbb{R}^N \to T^*\mathbb{R}^N$, defined by

$$((q_i)_{1\le i\le N}, (p_i)_{1\le i\le N}) \mapsto (-(q_{N-1}, \dots, q_1, q_N), -(p_{N-1}, \dots, p_1, p_N)). \quad (3.63)$$

Then S is a canonical linear involution satisfying $H_V \circ S = H_V$. Denote by Fix(S) the fixed point set of S. Then Fix(S) is the subset of all elements (q,p) in $T^*\mathbb{R}^N$ satisfying

$$(q_n, p_n) = -(q_{N-n}, p_{N-n}) \ \ \forall\, 1 \le n \le N-1 \text{ and } q_N = p_N = 0. \quad (3.64)$$

In particular, on Fix(S), $q_N = q_{N'+1} = 0$ and $p_N = p_{N'+1} = 0$. Note that on Fix(S), both the center of mass coordinate $Q = \frac{1}{N}\sum_{i=1}^N q_i$ and its momentum $P = \frac{1}{N}\sum_{i=1}^N p_i$ are identically 0. Hence Fix$(S) \subseteq \mathcal{M}$, where

$$\mathcal{M} := \{(q,p) \in T^*\mathbb{R}^N | Q = 0; P = 0\}.$$

We endow $\mathcal{M}$ with the symplectic structure induced from $T^*\mathbb{R}^N$.

The phase space of an FPU chain with N' moving particles satisfying Dirichlet boundary conditions is $T^*\mathbb{R}^{N'}$, endowed with the canonical symplectic structure $\sum_{i=1}^{N'} dq_i \wedge dp_i$. It can be embedded into $\mathcal{M}$ by the map $\Theta : T^*\mathbb{R}^{N'} \to \mathcal{M}$ defined by

$$(q_i, p_i)_{1\le i\le N'} \mapsto \frac{1}{\sqrt{2}}((q_i, p_i)_{1\le i\le N'}, (0,0), -(q_{N'-i}, p_{N'-i})_{0\le i\le N'-1}, (0,0)).$$

Note that $\Theta(T^*\mathbb{R}^{N'}) = Fix(S)$, i.e. Θ is a parametrization of Fix(S) and the pullback of the canonical symplectic form on $\mathcal{M}$ by Θ is $\sum_{i=1}^{N'} dq_i \wedge dp_i$, which means that Θ is canonical. It then follows that Fix(S) is a symplectic submanifold of $\mathcal{M}$.

We now express the equations defining Fix(S) locally near 0 as a subset of $\mathcal{M}$ in terms of the canonical coordinates $(x_k, y_k)_{1\le k\le N-1}$ provided by Theorem 1.3.1, or even more conveniently, in terms of the associated complex coordinates $(\zeta_k)_{1\le |k|\le N-1}$, defined for $1 \le k \le N-1$ as in (3.2) by

$$\begin{cases} \zeta_k = \frac{1}{\sqrt{2}}(x_k - i y_k), \\ \zeta_{-k} = \overline{\zeta_k} = \frac{1}{\sqrt{2}}(x_k + i y_k). \end{cases} \quad (3.65)$$

Denote as in (3.3) by $\mathcal{Z}$ the linear subspace of $\mathbb{C}^{2N-2}$ consisting of such vectors $(\zeta_k)_{1\le|k|\le N-1}$. In the sequel we also write $(\zeta_k)_{1\le k\le N-1}$ for the element $(\zeta_k)_{1\le|k|\le N-1} \in \mathcal{Z}$ and use the notations (in the case of s_n already introduced in (3.4))

$$c_n := \cos\frac{n\pi}{N}, \quad s_n := \sin\frac{n\pi}{N} \qquad (n \in \mathbb{Z}).$$

Define the map $S_{\mathcal{Z}} : \mathcal{Z} \to \mathcal{Z}$, given by

$$(\zeta_k)_{1\le k\le N-1} \mapsto (-e^{4\pi i k/N}\zeta_{N-k})_{1\le k\le N-1}. \quad (3.66)$$

Like the map $S : \mathcal{M} \to \mathcal{M}$, $S_{\mathcal{Z}}$ is a canonical linear involution. In fact, the maps S and $S_{\mathcal{Z}}$ are conjugate to each other under the coordinate change of Theorem 1.3.1. Before making this statement more precise, let us introduce a parametrization of the fixed point set $\mathrm{Fix}(S_{\mathcal{Z}})$ of the map $S_{\mathcal{Z}}$. Introduce

$$\mathcal{Z}_{Dir} := \{(\zeta_k)_{1\le|k|\le N'} \in \mathbb{C}^{2N'} \,|\, \overline{\zeta}_k = \zeta_{-k} \quad \forall\, 1 \le k \le N'\},$$

endowed with the *canonical* symplectic structure induced from $\mathbb{C}^{2N'}$, and the embedding $\Theta_{\mathcal{Z}} : \mathcal{Z}_{Dir} \to \mathcal{Z}$ mapping $(\zeta_k)_{1\le|k|\le N'}$ to the element $(\tilde{\zeta}_k)_{1\le k\le N} \in \mathcal{Z}$ given by

$$\frac{1}{\sqrt{2}}\left((\zeta_k)_{1\le k\le N'}, 0, (-e^{4\pi ik/N}\zeta_{N'+1-k})_{1\le k\le N'}\right).$$

Note that $\Theta_{\mathcal{Z}}(\mathcal{Z}_{Dir}) = Fix(S_{\mathcal{Z}})$, i.e. $\Theta_{\mathcal{Z}}$ is a parametrization of $\mathrm{Fix}(S_{\mathcal{Z}})$.

We prove the following lemma in Appendix C.

Lemma 3.3.1. *In terms of the complex variables $(\zeta_k)_{1\le|k|\le N-1}$ defined by Theorem 1.3.1, near 0, the map S is given by $S_{\mathcal{Z}}$. More precisely, if Ψ, defined near $0 \in \mathcal{Z}$, is the coordinate transformation given by Theorem 1.3.1, then $S \circ \Psi = \Psi \circ S_{\mathcal{Z}}$. In particular, locally near 0, the set $Fix(S_{\mathcal{Z}}) \subseteq \mathcal{Z}$, described by the equations*

$$e^{-2\pi ik/N}\zeta_k + e^{2\pi ik/N}\zeta_{N-k} = 0 \quad (1 \le k \le N-1), \tag{3.67}$$

is the image of the set $Fix(S)$ under Ψ^{-1}. Expressed in terms of the real variables $(x_k, y_k)_{1\le k\le N-1}$, the conditions (3.67) are given by

$$\begin{pmatrix} c_{2k} & -s_{2k} \\ s_{2k} & c_{2k} \end{pmatrix}\begin{pmatrix} x_k \\ y_k \end{pmatrix} + \begin{pmatrix} c_{2k} & s_{2k} \\ -s_{2k} & c_{2k} \end{pmatrix}\begin{pmatrix} x_{N-k} \\ y_{N-k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.68}$$

In particular, for $k = N' + 1\,(= N/2)$ we get $\zeta_{N'+1} = 0$ and therefore

$$(x_{N'+1}, y_{N'+1}) = (0, 0).$$

Corollary 3.3.2. *On $Fix(S_{\mathcal{Z}})$, for any $1 \le k \le \frac{N}{2}$,*

$$I_k = I_{N-k} \tag{3.69}$$

and

$$J_k J_{\frac{N}{2}-k} - M_k M_{\frac{N}{2}-k} = I_k I_{\frac{N}{2}-k}. \tag{3.70}$$

Moreover

$$I_{\frac{N}{2}} = 0. \tag{3.71}$$

Proof. In terms of the complex variables $(\zeta_k)_{1\le|k|\le N-1}$, $I_k = \zeta_k\zeta_{-k}$ for any $1 \le k \le N-1$. Hence on $\mathrm{Fix}(S_{\mathcal{Z}})$,

$$I_k = \zeta_k\zeta_{-k} = (-e^{4\pi ik/N}\zeta_{N-k})(-e^{-4\pi ik/N}\zeta_{-(N-k)}) = \zeta_{N-k}\zeta_{-(N-k)} = I_{N-k},$$

showing (3.69). The identity (3.71) follows from $\zeta_{\frac{N}{2}}|_{Fix(S_{\mathcal{Z}})} = 0$. To prove (3.70), we first conclude from (3.67) that on Fix($S_{\mathcal{Z}}$), for any $1 \leq k \leq N-1$,

$$\begin{aligned} J_k &= -c_{4k} I_k, \\ M_k &= -s_{4k} I_k. \end{aligned}$$

Hence on Fix($S_{\mathcal{Z}}$),

$$\begin{aligned} J_k J_{\frac{N}{2}-k} - M_k M_{\frac{N}{2}-k} &= I_k I_{\frac{N}{2}-k}(c_{4k} c_{4(\frac{N}{2}-k)} - s_{4k} s_{4(\frac{N}{2}-k)}) \\ &= I_k I_{\frac{N}{2}-k}(c_{4k}^2 + s_{4k}^2) \\ &= I_k I_{\frac{N}{2}-k}. \end{aligned}$$

This completes the proof of Corollary 3.3.2. □

From the definitions (1.9), (1.12), and (1.13) of the variables I_k, J_k, and M_k, and of the expressions R and $R_{\frac{N}{4}}$ one then obtains the following

Corollary 3.3.3. *On Fix($S_{\mathcal{Z}}$),*

$$R = 4 \sum_{1 \leq k < \frac{N}{4}} s_{2k} I_k I_{\frac{N}{2}-k} \quad \textit{and} \quad R_{\frac{N}{4}} = \begin{cases} I_{\frac{N}{4}}^2 & \textit{if } \frac{N}{4} \in \mathbb{N} \\ 0 & \textit{otherwise.} \end{cases}$$

It follows from Corollary 3.3.3 that on Fix($S_{\mathcal{Z}}$), the expression (1.14) is in Birkhoff normal form up to order four. This allows us to prove Theorem 1.4.1.

Proof of Theorem 1.4.1. We start with the resonant normal form (1.14) for even chains, $\frac{NP^2}{2} + H_{\alpha,\beta}(I) - R_{\alpha,\beta}(J,M) + O(|(x,y)|^5)$, where $H_{\alpha,\beta}(I)$ and $R_{\alpha,\beta}(J,M)$ are given by (1.7) and (1.11), respectively. Using the identity $I_k = I_{N-k}$, the terms in the decomposition (1.15) of $H_{\alpha,\beta}$, when restricted to Fix($S_{\mathcal{Z}}$), are given by

$$H^{(2)}(I) = 4 \sum_{k=1}^{N'} s_k \, I_k, \tag{3.72}$$

$$H^{(4)}_{\alpha,\beta}(I) = \frac{1}{N} \sum_{k=1}^{N'} d_k^+ I_k^2 + \frac{4(\beta - \alpha^2)}{2N} \sum_{\substack{1 \leq k,l \leq N' \\ k \neq l}} s_k s_l I_k I_l, \tag{3.73}$$

and

$$\frac{1}{2N} \sum_{k=1}^{\frac{N}{2}-1} d_k^- I_k I_{N-k} = \frac{1}{2N} \sum_{k=1}^{N'} d_k^- I_k^2. \tag{3.74}$$

From Corollary 3.3.3, we conclude that on Fix($S_{\mathcal{Z}}$),

$$\begin{aligned} -R_{\alpha,\beta}(J,M) &= -\frac{\beta - \alpha^2}{4N} \Big(R(J,M) + R_{\frac{N}{4}}(J,M) \Big) \\ &= -\frac{\beta - \alpha^2}{4N} \left(4 \sum_{1 \leq k < \frac{N}{4}} s_{2k} I_k I_{\frac{N}{2}-k} \underbrace{\; + \;\; I_{\frac{N}{4}}^2}_{\text{only if } \frac{N}{4} \in \mathbb{N}} \right). \end{aligned} \tag{3.75}$$

Formula (1.22) is then obtained by adding up (3.72)-(3.75), noting that $d_k^+ + \frac{d_k^-}{2} = \frac{1}{2}(\alpha^2 + 3(\beta - \alpha^2)s_k^2)$, and replacing I_k by its pullback $\frac{1}{2}I_k$ with respect to the parametrization $\Theta_{\mathcal{Z}}$ of $\mathrm{Fix}(S_{\mathcal{Z}})$ introduced above. $\square$

Chapter 4

Nondegeneracy and convexity

In this chapter we prove all claims on the nondegeneracy and convexity properties of the Hessians of the fourth-order Birkhoff normal forms of the odd periodic and Dirichlet chains, as stated in Theorems 1.2.3 and 1.4.3, respectively.

4.1 Odd periodic chains

We start with the Hessian $Q_{\alpha,\beta}$ of $H_{\alpha,\beta}(I)$ at $I = 0$, given by (3.41). In the process of investigating of $Q_{\alpha,\beta}$ we repeatedly encounter matrices of the form $E + \mathrm{diag}(\mu_1, \ldots, \mu_{N-1})$, where E is the $(N-1) \times (N-1)$-matrix

$$E := \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix} \tag{4.1}$$

and $(\mu_k)_{1 \le k \le N-1}$ are given complex numbers. The determinant of the matrix $E + \mathrm{diag}(\mu_1, \ldots, \mu_{N-1})$ can be explicitly computed.

Lemma 4.1.1. *Let $(\mu_k)_{1 \le k \le N-1}$ be given nonzero complex numbers. Then*

$$\det\left(E + \mathit{diag}(\mu_1, \ldots, \mu_{N-1})\right) = \left(1 + \sum_{k=1}^{N-1} \frac{1}{\mu_k}\right) \cdot \prod_{k=1}^{N-1} \mu_k. \tag{4.2}$$

In particular, $E + \mathit{diag}(\mu_1, \ldots, \mu_{N-1})$ is regular if and only if $\sum_{k=1}^{N-1} \frac{1}{\mu_k} \neq -1$.

Proof. Expanding $\det(E+\mathrm{diag}(\mu_1, \ldots, \mu_{N-1}))$ with respect to its rows it follows that

$$\det(E + \mathrm{diag}(\mu_1, \ldots, \mu_{N-1})) = \prod_{k=1}^{N-1} \mu_k + \sum_{k=1}^{N-1} \prod_{l \neq k} \mu_l.$$

This leads to formula (4.2). $\square$

First let us treat the β-chain, i.e. the case $\alpha = 0$, $\beta \neq 0$. The following proposition is relatd to earlier results of Rink [70].

Proposition 4.1.2. *Let N be odd and assume that $\alpha = 0$ in (1.4). Then the following holds:*

(i) The Birkhoff normal form of H_V up to order 4 is given by $\frac{NP^2}{2} + H_{0,\beta}(I)$ where

$$H_{0,\beta}(I) = 2\sum_{k=1}^{N-1} s_k I_k + \frac{\beta}{4N}\left(\sum_{k=1}^{N-1} s_k^2 I_k^2 + 2\sum_{\substack{l\neq m\\ 1\leq l,m\leq N-1}} s_l s_m I_l I_m\right). \tag{4.3}$$

(ii) For any $\beta \neq 0$, $H_{0,\beta}(I)$ is nondegenerate at $I = 0$.

Proof. The Birkhoff normal form (4.3) of H_V is given by the formula (3.41) evaluated at $\alpha = 0$. To investigate the Hessian $Q_{0,\beta}$ of $H_{0,\beta}(I)$ at $I = 0$ we write

$$Q_{0,\beta} = \frac{\beta}{4N}\,\Delta\, P\,\Delta, \tag{4.4}$$

where

$$\Delta = \operatorname{diag}\left(\sin\frac{k\pi}{N}\right)_{1\leq k\leq N-1} \tag{4.5}$$

and

$$P = 2\cdot\left(E - \frac{1}{2}\,\mathrm{Id}_{N-1}\right).$$

In view of (4.2) it follows that

$$\det Q_{0,\beta} = \left(\frac{\beta}{4N}\right)^{N-1}\cdot \det P \cdot \prod_{k=1}^{N-1}\sin^2\frac{k\pi}{N}$$

where by Lemma 4.1.1,

$$\det P = 2^{N-1}\,(1-2(N-1))\,(-1/2)^{N-1} = (-1)^N(2N-3) \neq 0.$$

Hence, if $\beta \neq 0$, $\det Q_{0,\beta} \neq 0$, and the nondegeneracy of $H_{0,\beta}(I)$ at $I = 0$ follows. □

Lemma 4.1.3. *If $\beta < 0$, then $Q_{0,\beta}$ has one negative eigenvalue, whereas if $\beta > 0$, then $Q_{0,\beta}$ has $N-2$ negative eigenvalues. In particular, for any $\beta \neq 0$, $Q_{0,\beta}$ is indefinite (and $H_{0,\beta}$ is therefore not convex).*

Proof. We want to use the decomposition (4.4) of $Q_{0,\beta}$ to show that $Q_{0,\beta}$ can be deformed continuously to $\frac{\beta}{4N}P$: Consider for $0 \leq t \leq 1$

$$Q_{0,\beta}(t) := \frac{\beta}{4N}(t\,\Delta + (1-t)\,\mathrm{Id})\;P\;(t\,\Delta + (1-t)\,\mathrm{Id}).$$

As $t\,\Delta + (1-t)\,\mathrm{Id}$ is positive definite for any $0 \le t \le 1$ and P is regular, $Q_{0,\beta}(t)$ is a symmetric regular $(N-1)\times(N-1)$-matrix for any $0 \le t \le 1$. For $t = 0$, $Q_{0,\beta}(0) = \frac{\beta}{4N}P$, whereas for $t = 1$, $Q_{0,\beta}(1) = Q_{0,\beta}$. Therefore, index$(Q_{0,\beta})$ (i.e. the number of negative eigenvalues of $Q_{0,\beta}$) coincides with index$\left(\frac{\beta}{4N}P\right)$. The eigenvalues of P are $\mu_1 = 2N-3$ with multiplicity one and $\mu_2 = -1$ with multiplicity $N-2$. □

We now turn to the case $\alpha \neq 0$.

Proposition 4.1.4. *Assume that N is odd and $\alpha \neq 0$ in (1.4). Then, for α fixed, $\det Q_{\alpha,\beta}$ is a polynomial in β of degree $N-1$ and has $N-1$ pairwise different real zeroes which we list in increasing order and denote by $\beta_k = \beta_k(\alpha)$ $(1 \le k \le N-1)$. They satisfy $0 < \beta_1 < \alpha^2$, $2\alpha^2 < \beta_2 < \ldots < \beta_{N-1}$ and contain the $(N-1)/2$ distinct numbers*

$$\alpha^2 \cdot \left(1 + \sin^{-2}\frac{k\pi}{N}\right) \quad \left(1 \le k \le \frac{N-1}{2}\right).$$

For these $\frac{N-1}{2}$ numbers, $Q_{\alpha,\beta}$ is not *isoenergetically nondegenerate.*

When considered as functions $\beta_k = \beta_k^{(N)}(\alpha)$ of N, the zeroes β_1 and β_2 satisfy

$$\beta_1 \to \alpha^2, \quad \beta_2 \to 2\alpha^2 \qquad (N \to \infty). \tag{4.6}$$

Moreover

$$\mathit{index}\,(Q_{\alpha,\beta}) = \begin{cases} 1 & \text{for } \beta < \beta_1, \\ 0 & \text{for } \beta_1 < \beta < \beta_2, \\ N-2 & \text{for } \beta > \beta_{N-1}. \end{cases}$$

Hence $H_{\alpha,\beta}$ is convex if and only if $\beta_1 < \beta < \beta_2$, in particular if $\frac{\beta}{\alpha^2} \in [1,2]$.

Proof. Since most of the statements of Propositions 4.1.4 are also true in the case where N is even, we do not assume a priori that N is odd, and we will mention explicitly when we make any assumption on the parity of N.

Fix $\alpha \in \mathbb{R} \setminus \{0\}$ and consider the map $\beta \mapsto \det(Q_{\alpha,\beta})$. It follows from (3.41) that $\det(Q_{\alpha,\beta})$ is a polynomial in β of degree at most $N-1$,

$$\det(Q_{\alpha,\beta}) = \sum_{j=0}^{N-1} c_j \beta^j,$$

where $c_0 = \det(Q_{\alpha,0})$ and $c_{N-1} = \det(Q_{0,1})$. By Proposition 4.1.2, $\det(Q_{0,1}) \neq 0$, hence the degree of the polynomial $\det(Q_{\alpha,\beta})$ is $N-1$. We claim that $\det(Q_{\alpha,\beta})$ has $N-1$ real zeroes (counted with multiplicities). For $|\beta|$ large enough, index$(Q_{\alpha,\beta})$ is equal to index$(Q_{0,\beta})$. By Lemma 4.1.3, index$(Q_{0,\beta})$ is $N-2$ for $\beta > 0$ and 1 for $\beta < 0$. Hence there exists $R > 0$ such that index$(Q_{\alpha,\beta}) = N-2$ for any $\beta > R$ and index$(Q_{\alpha,\beta}) = 1$ for any $\beta < -R$. For $\beta = \alpha^2$, the matrix Q_{α,α^2} is a positive multiple of the identity matrix, hence index$(Q_{\alpha,\alpha^2}) = 0$. It then follows that index$(Q_{\alpha,\beta})$ must change at least once in the open interval $(-\infty, \alpha^2)$ and at least $N-2$ times (counted with multiplicities) in (α^2, ∞). Since a change of index$(Q_{\alpha,\beta})$ induces a zero of $\det(Q_{\alpha,\beta})$ (counted with multiplicities), our consideration shows that $\beta \mapsto \det(Q_{\alpha,\beta})$ has $N-1$ real zeroes. Further we have $\beta_1(\alpha) < \alpha^2 < \beta_2(\alpha)$.

Next we prove that $\beta_1(\alpha) > 0$, i.e. that $Q_{\alpha,\beta}$ is regular for any $\beta \leq 0$. Write $Q_{\alpha,\beta}$ as a product,

$$Q_{\alpha,\beta} = \frac{\alpha^2 - \beta}{4N} \Delta \, P_{\alpha,\beta} \, \Delta, \tag{4.7}$$

with Δ defined by (4.5) and $P_{\alpha,\beta}$ by

$$P_{\alpha,\beta} = -2 \left(E + \operatorname{diag} \left(-\frac{1}{2} \left(1 + \frac{\gamma(\alpha,\beta)}{\sin^2 \frac{k\pi}{N}} \right) \right)_{1 \leq k \leq N-1} \right), \tag{4.8}$$

where E is given by (4.1) and

$$\gamma(\alpha,\beta) := \frac{\alpha^2}{\alpha^2 - \beta}. \tag{4.9}$$

As $-\infty < \beta \leq 0$ it follows that $0 < \gamma(\alpha,\beta) \leq 1$ and $-\frac{1}{2}\left(1 + \frac{\gamma(\alpha,\beta)}{\sin^2 \frac{k\pi}{N}}\right) < 0$ for any $1 \leq k \leq N-1$. Lemma 4.1.1 says that $P_{\alpha,\beta}$ is regular if $f(\gamma(\alpha,\beta)) \neq 0$ where

$$f(\gamma) := 1 - 2 \sum_{k=1}^{N-1} \left(1 + \gamma / \sin^2 \frac{k\pi}{N} \right)^{-1}. \tag{4.10}$$

Note that $f(\gamma)$ is increasing in $0 < \gamma \leq 1$ and $f(1)$ can be estimated as follows. Using that N is assumed to be odd one has

$$\begin{aligned} f(1) \, &= \, 1 - 4 \sum_{k=1}^{\frac{N-1}{2}} \frac{\sin^2 \frac{k\pi}{N}}{1 + \sin^2 \frac{k\pi}{N}} < 1 - 4 \frac{\sin^2 \frac{(N-1)\pi}{2N}}{1 + \sin^2 \frac{(N-1)\pi}{2N}} \\ &= 1 - 4 \frac{\cos^2 \frac{\pi}{2N}}{1 + \cos^2 \frac{\pi}{2N}} = -3 + \frac{4}{1 + \cos^2 \frac{\pi}{2N}}. \end{aligned}$$

As for $N \geq 3$

$$-3 + \frac{4}{1 + \cos^2 \frac{\pi}{2N}} < -3 + \frac{4}{1 + \cos^2 \frac{\pi}{6}} = -\frac{5}{7}$$

we conclude that $f(1) < 0$. Hence we have shown that $f(\gamma) < 0$ for $0 < \gamma \leq 1$, and therefore $P_{\alpha,\beta}$ is regular for $\beta \leq 0$ by Lemma 4.1.1. Hence we have proved that $0 < \beta_1(\alpha)$.

We now assume N to be odd. By writing $f(\gamma)$ as $f(\gamma) = 1 - 2 \sum_{k=1}^{N-1} \frac{s_k^2}{s_k^2 + \gamma}$ one sees that f is a rational function of $\gamma \in \mathbb{R}$ with poles of order one at $\gamma := -s_k^2$ $(1 \leq k \leq \frac{N-1}{2})$. For the derivative of f we obtain

$$f'(\gamma) = 2 \sum_{k=1}^{N-1} \frac{1}{(s_k^2 + \gamma)^2}. \tag{4.11}$$

Hence f is strictly increasing on all connected components of its domain. Further, for any of the poles $(-s_k^2)_{1\le k\le \frac{N-1}{2}}$ we have

$$f(\gamma)\nearrow\infty\quad\left(\gamma\nearrow -s_k^2\right)\quad\text{and}\quad f(\gamma)\searrow-\infty\quad\left(\gamma\searrow -s_k^2\right). \tag{4.12}$$

In addition, one sees that

$$f(\gamma)\nearrow 1\ \ (\gamma\to\infty)\quad\text{and}\quad f(\gamma)\searrow 1\ \ (\gamma\to-\infty). \tag{4.13}$$

It follows from (4.11)-(4.13) that f has precisely one zero in every of the $\frac{N-3}{2}$ bounded intervals $\left(-s_k^2,-s_{k-1}^2\right)$ $(1\le k\le\frac{N-3}{2})$ and (precisely) one zero γ_{N-1} in the unbounded interval $\left(-s_1^2,\infty\right)$ (since it is the largest zero of f, we denote this zero by γ_{N-1} even though it corresponds to β_1). The above analysis shows that $\gamma_{N-1}>0$ - we wil estimate γ_{N-1} more precisely below.

Next introduce $\mu_k:=-\frac{1}{2}(1+\gamma(\alpha,\beta)/s_k^2)$ and note that for β with $\gamma(\alpha,\beta)=-\sin^2\frac{k_0\pi}{N}$ for some $1\le k_0\le\frac{N-1}{2}$ (i.e. the poles of f mentioned above) one has $\mu_{k_0}=\mu_{N-k_0}=0$. As $k_0\ne N-k_0$ if $1\le k_0\le\frac{N-1}{2}$ it then follows that $P_{\alpha,\beta}$ has two equal rows and is therefore singular. Note that $\gamma(\alpha,\beta)=-s_{k_0}^2$ corresponds to $\frac{\beta}{\alpha^2}=1+s_{k_0}^{-2}$, and that $\gamma\in\left(-s_k^2,-s_{k-1}^2\right)$ corresponds to $\frac{\beta}{\alpha^2}\in\left(1+s_k^{-2},1+s_{k-1}^{-2}\right)$. Finally, as mentioned before, $\gamma_{N-1}>0$ corresponds to $\frac{\beta_1}{\alpha^2}<1$. Alltogether, we have proved that $\beta\mapsto\det(Q_{\alpha,\beta})$ has precisely $N-1$ pairwise different zeroes on $\mathbb{R}$. The statement about $\mathrm{index}(Q_{\alpha,\beta})$ easily follows from the above analysis.

In particular, $Q_{\alpha,\beta}$ is *not* nondegenerate for $\gamma(\alpha,\beta)=-s_{k_0}^2=-\sin^2\frac{k_0\pi}{N}$. It remains to show the statement that $Q_{\alpha,\beta}$ is also not *isoenergetically* nondegenerate for these parameter values. We thus assume that $\gamma(\alpha,\beta)=-\sin^2\frac{k_0\pi}{N}$ for some fixed $1\le k_0\le\frac{N-1}{2}$. Instead of directly disproving the condition (2.7) for isoenergetic nondegeneracy, we show the stronger statement that the $(N-1)\times N$-matrix (2.8), in this case given by the $(N-1)\times N$-matrix

$$\tilde{Q}_{\alpha,\beta}:=\left(\begin{array}{ccccc|c}
s_1-\gamma & & & \vdots & & d_\gamma^{(N)}\cdot s_1\\
 & s_2-\gamma & & \cdots\ (2s_ls_m)_{l<m} & \cdots & \vdots\\
 & & \ddots & \vdots & & \\
 & \vdots & & \ddots & & \\
\cdots & (2s_ls_m)_{l>m} & \cdots & s_{N-2}-\gamma & & \vdots\\
 & \vdots & & & s_{N-1}-\gamma & d_\gamma^{(N)}\cdot s_{N-1}
\end{array}\right)$$

with $d_\gamma^{(N)}=\frac{8N}{\beta-\alpha^2}$, does not have rank $N-1$. As explained in section 2.4, this then implies that the isoenergeticity condition (2.7) is violated. We argue indirectly and assume that the matrix $\tilde{Q}_{\alpha,\beta}$ has rank $N-1$, which is equivalent to the regularity of at least one of its N square submatrices of size $N-1$. Note that we can replace the nonzero number $d_\gamma^{(N)}$ by another nonzero number, namely 2, without changing the rank of $\tilde{Q}_{\alpha,\beta}$ - this simplifies the following computations. Since $Q_{\alpha,\beta}$ is singular, it remains to consider the $N-1$ submatrices of $\tilde{Q}_{\alpha,\beta}$ given by $\tilde{Q}_{\alpha,\beta}$ without its last column and its l-th column replaced by the vector $2s_l(s_1,\ldots,s_{N-1})$ (the additional multiplication of the l-th column by the nonzero factor s_l again simplifies the following computations); we denote these $N-1$ matrices by $Q^{(l)}_{\alpha,\beta}$ $(1 \le l \le N-1)$. By the same procedure as above, we decompose $Q^{(l)}_{\alpha,\beta}$ as

$$Q^{(l)}_{\alpha,\beta} = \Delta\, P^{(l)}_{\alpha,\beta}\, \Delta,$$

where Δ is again given by (4.5) and $P^{(l)}_{\alpha,\beta}$ by $P^{(l)}_{\alpha,\beta} = -2(E + \operatorname{diag}(p^{(l)}_{\alpha,\beta})_{1\le k\le N-1})$ with

$$(p^{(l)}_{\alpha,\beta})_k = \begin{cases} -\frac{1}{2}\left(1 + \frac{\gamma(\alpha,\beta)}{\sin^2\frac{k\pi}{N}}\right) & (k \neq l) \\ 0 & (k = l). \end{cases}$$

It follows from $\gamma(\alpha,\beta) = -s^2_{k_0}$ that for any $1 \le l \le N-1$, $(p^{(l)}_{\alpha,\beta})_{k_0} = (p^{(l)}_{\alpha,\beta})_{N-k_0} = 0$. Thus, for any $1 \le l \le N-1$, $P^{(l)}_{\alpha,\beta}$ has two equal rows and is therefore singular. As explained above, this implies that $Q_{\alpha,\beta}$ is not isoenergetically nondegenerate. Note that we do not check Rüssmann's higher order nondegeneracy conditions, so that it remains an open question whether some variant of the KAM theorem can be applied to the odd periodic FPU chain for these exceptional β's.

We now turn to the asymptotic statements (4.6). It follows from the above analysis that

$$\beta_2^{(N)} = \alpha^2\left(1 + \sin^{-2}\frac{(N-1)/2 * \pi}{N}\right) = \alpha^2\left(1 + \sin^{-2}\left(\frac{\pi}{2} - \frac{\pi}{2N}\right)\right),$$

and one sees that $\beta_2^{(N)} \to 2\alpha^2$ for $N \to \infty$. On the other hand, proving that $\beta_1^{(N)} \to \alpha^2$ for $N \to \infty$ turns out to be considerably more difficult; nevertheless, we consider it justified to prove this in detail since the two asymptotic statements together give us precise information on the length of the "interval of convexity" of $Q_{\alpha,\beta}$ in the limit $N \to \infty$.

We mentioned above that $\beta \to \det Q_{\alpha,\beta}$ has $N-1$ pairwise distinct zeros, of which $\frac{N-1}{2}$ ones are in terms of $\gamma \equiv \gamma(\alpha,\beta)$ given by $\gamma = -\sin^2\frac{k\pi}{N}$, $1 \le k \le \frac{N-1}{2}$, and the other $\frac{N-1}{2}$ ones are zeroes of the meromorphic function f defined in (4.10), $f(\gamma) = 1 - 2\sum_{k=1}^{N-1}\left(1 + \gamma/\sin^2\frac{k\pi}{N}\right)^{-1}$, whose $\frac{N-1}{2}$ poles of order 1 are exactly the other zeros of $\beta \to \det Q_{\alpha,\beta}$ mentioned before.

Thus, if we multiply $f(\gamma)$ by $\prod_{k=1}^{N-1}(\gamma + s_k^2) = \left(\prod_{k=1}^{(N-1)/2}(\gamma + s_k^2)\right)^2$, we obtain a polynomial p whose $N-1$ zeros are precisely the $N-1$ zeros of $\det Q_{\alpha,\beta}$. (The zeroes of f are also zeroes of p, and the first-order poles of f are first-order zeroes of p.) Explicitly, $p(\cdot)$ is then given by

$$p(\gamma) = \prod_{k=1}^{N-1}(\gamma + s_k^2) - 2\sum_{k=1}^{N-1} s_k^2 \prod_{\substack{1\le l\le N-1 \\ l\neq k}}(\gamma + s_l^2). \tag{4.14}$$

Note that when $p(\gamma)$ is ordered by powers of γ, the coefficients are *symmetric* polynomials of the $N-1$ variables $(s_k)_{1\leq k\leq N-1}$, i.e. we can express these coefficients through the N basic symmetric polynomials $(\Pi_n(t_1,\ldots,t_{N-1}))_{0\leq n\leq N-1}$ evaluated for $t_k := s_k^2$. These basic symmetric polynomials are given by

$$\Pi_0 := 1, \tag{4.15}$$

$$\Pi_n(t_1,\ldots,t_{N-1}) := \sum_{1\leq i_1<\ldots<i_n\leq N-1} t_{i_1}\cdot\ldots\cdot t_{i_n} \ (1\leq n\leq N-1). \tag{4.16}$$

Ordering $p(\gamma)$ by powers of γ, this method yields

$$p(\gamma) = \sum_{r=0}^{N-1} \Pi_r(s_1^2,\ldots,s_{N-1}^2)(1-2r)\gamma^{N-1-r}. \tag{4.17}$$

We now evaluate the polynomials $(\Pi_n)_{0\leq n\leq N-1}$ given by (4.15) and (4.16) for $t_k = s_k^2$. We give the (rather technical) proof of the following lemma in Appendix D.

Lemma 4.1.5. *For any* $0\leq r\leq N-1$,

$$\Pi_r\left(s_1^2,\ldots,s_{N-1}^2\right) = 4^{-r}\frac{N}{N-r}\binom{2N-r-1}{r}. \tag{4.18}$$

Substituting (4.18) into (4.17), we obtain

$$\begin{aligned} p(\gamma) &= \sum_{r=0}^{N-1}(1-2r)\Pi_r(s_1^2,\ldots,s_{N-1}^2)\gamma^{N-1-r} \\ &= 4^{-(N-1)}\sum_{r=0}^{N-1}(1-2r)\frac{N}{N-r}\binom{2N-r-1}{r}(4\gamma)^{N-1-r}. \end{aligned}$$

The index substitution $s = N-1-r$ then leads to (note that $1-2r = 3-2(N-s)$)

$$p(\gamma) = 4^{-(N-1)}\sum_{s=0}^{N-1}(3-2(N-s))\frac{N}{s+1}\binom{N+s}{N-1-s}(4\gamma)^s.$$

We now omit the factor $4^{-(N-1)}$ (since it does not influence the zeros of p) and set $x := 4\gamma$, which leads to the polynomial

$$P(x) := \sum_{s=0}^{N-1}\frac{(3-2(N-s))N}{s+1}\binom{N+s}{N-1-s}x^s = x^{N-1} - 2N\cdot x^{N-2} + O(x^{N-3}).$$

By Vieta's theorem (see e.g. [90]), the $N-1$ zeros $x_1,\ldots,x_{N-1}$ of P must satisfy

$$\sum_{k=1}^{N-1} x_k = 2N. \tag{4.19}$$

Recall from above that we already know precisely $\frac{N-1}{2}$ zeroes, namely $x_k = -4\sin^2\frac{k\pi}{N}$ ($1 \le k \le \frac{N-1}{2}$), and from $\frac{N-3}{2}$ zeros we know that they satisfy $-4\sin^2\frac{k\pi}{N} < x_k < -4\sin^2\frac{(k+1)\pi}{N}$ ($\frac{N+1}{2} \le k \le N-2$). Hence, by (4.19), the remaining zero x_{N-1} is located in the open interval

$$\left(2N + 4\sum_{k=1}^{\frac{N-1}{2}} \sin^2\frac{k\pi}{N} + 4\sum_{k=\frac{N+3}{2}}^{N-1} \sin^2\frac{k\pi}{N}\,,\ 2N + 4\sum_{k=1}^{\frac{N-1}{2}} \sin^2\frac{k\pi}{N} + 4\sum_{k=\frac{N+1}{2}}^{N-2} \sin^2\frac{k\pi}{N}\right)$$
$$= \left(2N + 4\cdot\left(\frac{N}{2} - \sin^2\frac{(N-1)\pi}{2N}\right),\ 2N + 4\cdot\left(\frac{N}{2} - \sin^2\frac{\pi}{N}\right)\right)$$
$$= \left(4N - 4\sin^2\left(\frac{\pi}{2}\left(1-\frac{1}{N}\right)\right),\ 4N - 4\sin^2\frac{\pi}{N}\right) \subset (4(N-1),\ 4N).$$

In terms of $\gamma = x/4$, we conclude that the zero γ_{N-1} satisfies

$$N - 1 < \gamma_{N-1} < N. \tag{4.20}$$

Numerical evidence (Mathematica computations) actually suggests that we have the asymptotic formula $\gamma_{N-1} = N - \frac{3}{4} + o(1)$ ($N \to \infty$), in accordance with the analytically derived estimate (4.20). In terms of $\beta = \alpha^2(1-\frac{1}{\gamma})$, it follows that the corresponding zero β_1 can be estimated by

$$\alpha^2\left(1 - \frac{1}{N-1}\right) < \beta_1 < \alpha^2\left(1-\frac{1}{N}\right),$$

in particular we have $\beta_1 \to \alpha^2$ for $N \to \infty$, as claimed. □

Proof of Theorem 1.2.3. Part (i) is proved by Proposition 4.1.4, whereas (ii) follows from Proposition 4.1.2 and Lemma 4.1.3. □

4.2 Dirichlet chains

We now turn to the chains with Dirichlet boundary conditions. Again, we first consider the case $\alpha = 0$. Note that the numbers $(s_k)_{1\le k\le N'}$ are as before (cf. (3.4)) defined by $s_k = \sin\frac{k\pi}{N}$, which however should now be read as $s_k = \sin\frac{k\pi}{2N'+2}$ (recall that $N = 2N'+2$).

Proposition 4.2.1. *Assume that $\alpha = 0$ in (1.4). Then the following holds:*

(i) The Birkhoff normal form of H_V with Dirichlet boundary conditions up to order 4 is given by $\frac{(N'+1)P^2}{2} + H^D_{0,\beta}(I)$ where

$$H^D_{0,\beta}(I) = 2\sum_{k=1}^{N'} s_k I_k + \frac{\beta}{16(N'+1)}\left(\sum_{k=1}^{N'} 3s_k^2 I_k^2 \quad \underbrace{+\frac{1}{2}I^2_{\frac{N'+1}{2}}}_{only\ if\ \frac{N'+1}{2}\in\mathbb{N}}\right.$$
$$\left. +4\sum_{\substack{l\ne m\\ 1\le l,m\le N'}} s_l s_m I_l I_m - \sum_{k=1}^{N'} s_{2k} I_k I_{N'+1-k}\right). \tag{4.21}$$

(ii) For any $\beta \neq 0$, $H^D_{0,\beta}(I)$ is nondegenerate at $I = 0$.

Proof. The Birkhoff normal form (4.21) of H_V with Dirichlet boundary conditions is given by the formula (1.22) evaluated at $\alpha = 0$. To investigate the Hessian of $Q^D_{0,\beta}$ of $H^D_{0,\beta}(I)$ at $I = 0$, we write

$$Q^D_{0,\beta} = \frac{2\beta}{16(N'+1)} \Delta^{N'} P^D \Delta^{N'}, \tag{4.22}$$

where $\Delta^{N'} = \operatorname{diag}\left(\sin \frac{k\pi}{2N'+2}\right)_{1 \le k \le N'}$ and P^D is the $N' \times N'$-matrix which for N' even resp. odd is of the form

$$\underbrace{\begin{pmatrix} 3 & 4 & \dots & & & & & \dots & 4 & 2 \\ 4 & 3 & 4 & \dots & & & \dots & 4 & 2 & 4 \\ \vdots & & \ddots & & & & & \ddots & & \vdots \\ & & & 3 & 4 & 4 & 2 & & & \\ & & & 4 & 3 & 2 & 4 & & & \\ & & & 4 & 2 & 3 & 4 & & & \\ & & & 2 & 4 & 4 & 3 & & & \\ \vdots & & \ddots & & & & & \ddots & & \vdots \\ 4 & 2 & 4 & \dots & & & \dots & 4 & 3 & 4 \\ 2 & 4 & \dots & & & & & \dots & 4 & 3 \end{pmatrix}}_{(N' \text{ even})}, \quad \underbrace{\begin{pmatrix} 3 & 4 & \dots & & & & \dots & 4 & 2 \\ 4 & 3 & 4 & \dots & & \dots & 4 & 2 & 4 \\ \vdots & & \ddots & & & & \ddots & & \vdots \\ & & & 3 & 4 & 2 & & & \\ & & & 4 & 2 & 4 & & & \\ & & & 2 & 4 & 3 & & & \\ \vdots & & \ddots & & & & \ddots & & \vdots \\ 4 & 2 & 4 & \dots & & \dots & 4 & 3 & 4 \\ 2 & 4 & \dots & & & & \dots & 4 & 3 \end{pmatrix}}_{(N' \text{ odd})},$$

where we used that $s_{2k} = 2s_k c_k = 2s_k s_{N'+1-k}$ and, if $\frac{N'+1}{2} \in \mathbb{N}$, $s^2_{\frac{N'+1}{2}} = \frac{1}{2}$. It follows that

$$\det Q^D_{0,\beta} = \left(\frac{2\beta}{16(N'+1)}\right)^{N-1} \cdot \det P^D \cdot \prod_{k=1}^{N'} \sin^2 \frac{k\pi}{2N'+2}.$$

In order to see that P^D is nonsingular, observe that $\det P^D \in \mathbb{Z}$. For N' even we show that $\det P^D \equiv 1 \bmod 2$. Note that in this case the diagonal of P^D consists of 3's only. Therefore $\det P^D \equiv 3^{N'} \bmod 2 \equiv 1 \bmod 2$. If N' is odd, the same argument shows that $\det P \equiv 2 \bmod 4$. Hence, if $\beta \neq 0$, $\det Q^D_{0,\beta} \neq 0$, and the nondegeneracy of the Hessian of $H^D_{0,\beta}(I)$ at $I = 0$ follows. $\square$

Lemma 4.2.2. *If $\beta < 0$, then $Q^D_{0,\beta}$ has $\lceil \frac{N'+1}{2} \rceil$ negative eigenvalues, whereas if $\beta > 0$, then $Q^D_{0,\beta}$ has $\lfloor \frac{N'-1}{2} \rfloor$ negative eigenvalues. In particular, for any $\beta \neq 0$, $Q^D_{0,\beta}$ is indefinite (and $H^D_{0,\beta}$ is therefore not convex).*

Proof. We want to use the decomposition (4.22) of $Q^D_{0,\beta}$ to show that $Q^D_{0,\beta}$ can be deformed continuously to $\frac{2\beta}{16(N'+1)} P^D$: Consider for $0 \le t \le 1$

$$Q^D_{0,\beta}(t) := \frac{2\beta}{16(N'+1)} (t\,\Delta^{N'} + (1-t)\,\mathrm{Id})\; P^D\; (t\,\Delta^{N'} + (1-t)\,\mathrm{Id}).$$

As $t\,\Delta^{N'} + (1-t)\,\mathrm{Id}$ is positive definite for any $0 \le t \le 1$ and P^D is regular and symmetric, $Q^D_{0,\beta}(t)$ is a symmetric regular $N' \times N'$-matrix for any $0 \le t \le 1$. For $t = 0$, $Q^D_{0,\beta}(0) = \frac{2\beta}{16(N'+1)} P^D$, whereas for $t = 1$, $Q^D_{0,\beta}(1) = Q^D_{0,\beta}$. Therefore, index$(Q^D_{0,\beta})$ (i.e. the number of negative eigenvalues of $Q^D_{0,\beta}$) coincides with index$(\frac{2\beta}{16(N'+1)} P^D)$. To list the eigenvalues of P^D, we distinguish between N' even and odd.

If N' is even, the eigenvalues of P^D are $4N' - 3$ (with multiplicity one), 1 (with multiplicity $\frac{N'}{2}$), and -3 (with multiplicity $\frac{N'}{2} - 1$), hence P^D has $\frac{N'}{2} - 1$ negative eigenvalues. If N' is odd, the eigenvalues of P^D are 1 (with multiplicity $\frac{N'-1}{2}$), -3 (with multiplicity $\frac{N'-3}{2}$), and $\frac{1}{2}(4N'-5)(1 \pm \sqrt{1 + \frac{8(4N'-1)}{(4N'-5)^2}})$ (each with multiplicity one), hence P^D has $\frac{N'-1}{2}$ negative eigenvalues. These facts are verified in Appendix E. The claim of the lemma now follows immediately. □

We now turn to the case $\alpha \neq 0$.

Proposition 4.2.3. *Assume that $\alpha \neq 0$ in (1.4). Then, for α fixed, $\det Q^D_{\alpha,\beta}$ is a polynomial in β of degree N' and has N' real zeroes (counted with multiplicities). When denoted by $\beta_k = \beta_k(\alpha)$ $(1 \le k \le N')$ and listed in increasing order, they satisfy*

$$\beta_1 \le \ldots \le \beta_{\lceil \frac{N'+1}{2} \rceil} < \alpha^2 < \beta_{\lceil \frac{N'+3}{2} \rceil} \le \ldots \le \beta_{N'}.$$

Moreover

$$\mathit{index}\,(Q^D_{\alpha,\beta}) = \begin{cases} \lceil \frac{N'+1}{2} \rceil & \text{for } \beta < \beta_1 \\ 0 & \text{for } \beta_{\lceil \frac{N'+1}{2} \rceil} < \beta < \beta_{\lceil \frac{N'+3}{2} \rceil} \\ \lfloor \frac{N'-1}{2} \rfloor & \text{for } \beta > \beta_{N'} \end{cases}$$

Hence $H^D_{\alpha,\beta}$ is convex if and only if $\beta_{\lceil \frac{N'+1}{2} \rceil} < \beta < \beta_{\lceil \frac{N'+3}{2} \rceil}$.

Proof. Fix $\alpha \in \mathbb{R} \setminus \{0\}$ and consider the map $\beta \mapsto \det(Q^D_{\alpha,\beta})$. It follows from (1.22) that $\det(Q^D_{\alpha,\beta})$ is a polynomial in β of degree at most N',

$$\det(Q^D_{\alpha,\beta}) = \sum_{j=0}^{N'} r_j \beta^j,$$

where $r_0 = \det(Q^D_{\alpha,0})$ and $r_{N'} = \det(Q^D_{0,1})$. By Proposition 4.2.1, $\det(Q^D_{0,1}) \neq 0$, hence the degree of the polynomial $\det(Q^D_{\alpha,\beta})$ is N'. We claim that $\det(Q^D_{\alpha,\beta})$ has N' real zeroes (counted with multiplicities). For $|\beta|$ large enough, index$(Q^D_{\alpha,\beta})$ is equal to index$(Q^D_{0,\beta})$. By Lemma 4.2.2, index$(Q^D_{0,\beta})$ is $\lfloor \frac{N'-1}{2} \rfloor$ for $\beta > 0$ and $\lceil \frac{N'+1}{2} \rceil$ for $\beta < 0$. Hence there exists $R > 0$ such that index$(Q^D_{\alpha,\beta}) = \lfloor \frac{N'-1}{2} \rfloor$ for any $\beta > R$ and index$(Q^D_{\alpha,\beta}) = \lceil \frac{N'+1}{2} \rceil$ for any $\beta < -R$. For $\beta = \alpha^2$, Q^D_{α,α^2} is a positive multiple of the identity matrix, hence index$(Q^D_{\alpha,\alpha^2}) = 0$. It then follows that, when counted with multiplicities, $\mathrm{index}(Q^D_{\alpha,\beta})$ must change at least $\lceil \frac{N'+1}{2} \rceil$ times in the open interval $(-\infty, \alpha^2)$ and at least $\lfloor \frac{N'-1}{2} \rfloor$ times in (α^2, ∞). Since a change of $\mathrm{index}(Q^D_{\alpha,\beta})$ induces a real zero of $\det(Q^D_{\alpha,\beta})$, our consideration shows that $\beta \mapsto \det(Q^D_{\alpha,\beta})$ has N' real zeroes. Further we have $\beta_{\lceil \frac{N'+1}{2} \rceil}(\alpha) < \alpha^2 < \beta_{\lceil \frac{N'+3}{2} \rceil}(\alpha)$. □

Proof of Theorem 1.4.3. Part (i) is proved by Proposition 4.2.3, whereas (ii) follows from Propostion 4.2.1 and Lemma 4.2.2. □

Chapter 5

The foliation of the phase space by the moment map of an integrable approximation of even periodic chains

In this chapter we describe the geometry of the moment map of the truncated resonant normal form (1.14) for any even periodic FPU chain H_V with potential V whose expansion (1.4) satisfies $(\alpha, \beta) \neq (0, 0)$. The case $\beta = \alpha^2$ is special as in this case the normal form (1.14) is the Birkhoff normal form of order four of the Toda lattice. Its foliation is well known - it is the one of uncoupled harmonic oscillators. Hence we will concentrate on the case $\beta \neq \alpha^2$ only. The special case $\alpha = 0$ has been partially studied by Rink [71]. Surprisingly, it turns out that many of his results continue to hold in the general case. Using the notation $\tilde{k} \equiv \tilde{k}(k) = \frac{N}{2} - k$, the integrals of Theorem 1.3.2 can be grouped as follows:

$$(\mathcal{H}_k, \mathcal{H}_{\tilde{k}}, L_k, K_k)_{1 \leq k < \frac{N}{4}}, \qquad I_{\frac{N}{2}}, \qquad \mathcal{H}_{\frac{N}{4}}, K_{\frac{N}{4}}, \tag{5.1}$$

where for $1 \leq k < \frac{N}{2}$

$$\mathcal{H}_k := I_k + I_{N-k}, \quad L_k := I_k - I_{\tilde{k}}.$$

(Here we used that $L_k = (I_k + I_{\frac{N}{2}+k}) - (I_{\frac{N}{2}-k} + I_{\frac{N}{2}+k})$ is the difference of two integrals listed in Theorem 1.3.2.)

Using the assumption $\alpha^2 - \beta \neq 0$, we rewrite the integrals $(K_l)_{1 \leq l \leq \frac{N}{4}}$ as follows. We again use the bifurcation parameter

$$\gamma \equiv \gamma(\alpha, \beta) := \frac{\alpha^2}{\alpha^2 - \beta} \tag{5.2}$$

and note that

$$d_k^- = -\alpha^2 + (\beta - \alpha^2)s_k^2 = (\beta - \alpha^2)(\gamma + s_k^2).$$

For any $1 \leq k < \frac{N}{4}$, one has by (1.17)-(1.19) (note that $s_{\tilde{k}} = c_k$ for any $1 \leq k < \frac{N}{4}$)

$$\begin{aligned} K_k &= -d_k^-(J_k^2 + M_k^2) - d_{\tilde{k}}^-(J_{\tilde{k}}^2 + M_{\tilde{k}}^2) + 2(\beta - \alpha^2)s_{2k}(J_k J_{\tilde{k}} - M_k M_{\tilde{k}}) \\ &= s_{2k}(\alpha^2 - \beta)\left(\frac{\gamma + s_k^2}{s_{2k}}(M_k^2 + J_k^2) + \frac{\gamma + c_k^2}{s_{2k}}(M_{\tilde{k}}^2 + J_{\tilde{k}}^2) + 2(M_k M_{\tilde{k}} - J_k J_{\tilde{k}})\right), \end{aligned}$$

whereas for $k = \frac{N}{4}$,

$$K_k = \alpha^2 J_k^2 - (\beta - 2\alpha^2)M_k^2 = (\alpha^2 - \beta)\left(\gamma J_k^2 + (1 + \gamma)M_k^2\right).$$

In the sequel, for simplicity we will *omit* the factor $s_{2k}(\alpha^2 - \beta)$ in K_k, since it does not influence the geometry of the level sets of the integrals $(K_k)_{1 \leq k \leq \frac{N}{4}}$.

Each of the $\llcorner \frac{N}{4} \lrcorner + 1$ groups of integrals listed in (5.1) depends only on a subset of the variables $\{(x_k, y_k)_{1 \leq k \leq N-1}\}$. These subsets form a disjoint partition of $\{(x_k, y_k)_{1 \leq k \leq N-1}\}$. More precisely, the following result holds.

Proposition 5.0.4. *The phase space $T^*\mathbb{R}^{N-1}$ of the truncated resonant normal form H_V^{trunc} given by (1.14) is the direct sum of invariant symplectic subspaces*

$$T^*\mathbb{R}^{N-1} = \bigoplus_{0 \leq k \leq \frac{N}{4}} \mathcal{P}_k,$$

where

$$\mathcal{P}_k = \{(x_j, y_j)_{1 \leq j \leq N-1} \in T^*\mathbb{R}^{N-1} | x_j = y_j = 0 \;\forall\, j \notin \{k, N-k, \tilde{k}, N-\tilde{k}\}\}.$$

The foliation of $T^*\mathbb{R}^{N-1}$ by level sets of the integrals (5.1) is the Cartesian product of the foliations of the $\mathcal{P}_k$.

We now analyze the foliations of $\mathcal{P}_0$, $\mathcal{P}_{\frac{N}{4}}$, and $\mathcal{P}_k$ for $0 < k < \frac{N}{4}$ separately. Note that $\mathcal{P}_0$ and $\mathcal{P}_k$ $(0 < k < \frac{N}{4})$ can be canonically identified with $T^*\mathbb{R}$ and $T^*\mathbb{R}^4$, respectively, and $\mathcal{P}_{\frac{N}{4}}$ with $T^*\mathbb{R}^2$ (if $\frac{N}{4} \in \mathbb{N}$) or $\{0\}$ (if $\frac{N}{4} \notin \mathbb{N}$).

5.1 Foliation of $\mathcal{P}_0$

One easily sees that $I_{\frac{N}{2}}$ foliates $T^*\mathbb{R}$ by circles, centered at the origin.

5.2 Foliation of $\mathcal{P}_{\frac{N}{4}}$ for $\frac{N}{4} \in \mathbb{Z}$

Let us study the geometry of the moment map $\mathcal{M} : T^*\mathbb{R}^2 \to \mathbb{R}^2$ defined by the integrable system with commuting integrals $H \equiv H_{\frac{N}{4}}$ and $K \equiv K_{\frac{N}{4}}$. It is convenient to introduce the following notation. Denote the standard coordinates of $T^*\mathbb{R}^2$ by $(x, y) = (x_1, x_2, y_1, y_2)$ and introduce the action variables $I_j = \frac{1}{2}(x_j^2 + y_j^2)$ $(j = 1, 2)$, as well as the Hopf variables M, J, L given as in (1.9),

$$(M, J, L) = \frac{1}{2}(x_1y_2 - x_2y_1, x_1x_2 + y_1y_2, I_1 - I_2).$$

Then the moment map $\mathcal{M} = (H, K)$ takes the form

$$H = \frac{1}{2}(I_1 + I_2) \quad \text{and} \quad K = (1 + \gamma)M^2 + \gamma J^2.$$

As already remarked in (1.10) one has

$$M^2 + J^2 + L^2 = H^2.$$

Further, we may replace K by K_γ given by

$$K_\gamma := \begin{cases} (1 + \gamma)M^2 + \gamma J^2 & \gamma \notin \{-1, 0\}, \\ M & \gamma = 0, \\ J & \gamma = -1. \end{cases}$$

First observe that the origin $(x, y) = (0, 0)$ of $T^*\mathbb{R}^2$ is the only critical point of $\mathcal{M}$ with rank $d_{(x,y)}\mathcal{M} = 0$. Moreover,

$$\mathcal{M}^{-1}\{(0, 0)\} = \{(0, 0)\}.$$

The critical points $(x, y) \in T^*\mathbb{R}^2 \setminus \{(0, 0)\}$ with rank $d_{(x,y)}\mathcal{M} = 1$ are analyzed by symplectic reduction via the Hamiltonian vector field of H. On the sphere $\mathbb{S}^3_\rho = \{H = \rho^2/4\}$ of radius $\rho > 0$ in $T^*\mathbb{R}^2$ define the Hopf map

$$\mathcal{F} : \mathbb{S}^3_\rho \to \mathbb{S}^2_r, (x, y) \mapsto (M, J, L)$$

where $r = \sqrt{M^2 + J^2 + L^2}|_{\mathbb{S}^3_\rho} = H|_{\mathbb{S}^3_\rho} = \frac{\rho^2}{4}$. The fibers of $\mathcal{F}$ are circles obtained by the $\mathbb{S}^1$-action of H. The reduced system is then given by $(\mathbb{S}^2_r, X_\gamma)$ where X_γ denotes the reduced Hamiltonian vector field induced by K_γ. To compute X_γ, note that the equations of motion in the reduced system corresponding to the Hamiltonian K_γ are given by

$$\frac{d}{dt}\begin{pmatrix} M \\ J \\ L \end{pmatrix} = \begin{pmatrix} M \\ J \\ L \end{pmatrix} \times \begin{pmatrix} \partial_M K_\gamma \\ \partial_J K_\gamma \\ \partial_L K_\gamma \end{pmatrix}. \tag{5.3}$$

Indeed, following the procedure of reduction in section I.5 of [15], formula (5.3) follows from

$$\dot{w}_j = \{w_j, K_\gamma\} = \sum_{i=1}^{3} \partial_{w_i} K_\gamma \{w_j, w_i\} \qquad (1 \leq j \leq 3)$$

and the commutation relations of the variables $(w_1, w_2, w_3) = (M, J, L)$ given by Lemma 3.2.3. We then obtain

$$X_\gamma = \begin{cases} (-2\gamma JL, 2(1+\gamma)ML, -2MJ) & \gamma \notin \{-1, 0\}, \\ (0, L, -J) & \gamma = 0, \\ (-L, 0, M) & \gamma = -1. \end{cases}$$

It turns out that the foliation of $\mathbb{S}^2_r$ by level sets of K_γ depends on the bifurcation parameter γ. If $\gamma = 0$, then $(\pm r, 0, 0)$ are the only two fixed points of X_0. They are both elliptic and the level sets of K_0 in $\mathbb{S}^2_r \setminus \{(\pm r, 0, 0)\}$ are circles. Similarly, for $\gamma = -1$, $(0, \pm r, 0)$ are the only two fixed points of X_{-1}. They are both elliptic and the level sets of K_{-1} in $\mathbb{S}^2_r \setminus \{(0, \pm r, 0)\}$ are circles. Now let us consider the case $\gamma \notin \{-1, 0\}$. Then X_γ admits six fixed points,

$$(\pm r, 0, 0), \quad (0, \pm r, 0), \quad (0, 0, \pm r)$$

where two of them are hyperbolic and the remaining four elliptic. Note that the corresponding values of K_γ are $(1+\gamma)r^2$, γr^2, and 0, respectively, and that the two hyperbolic fixed points are contained in the same connected component of the inverse image of K_γ in $\mathbb{S}^2_r$. This component consists of two great circles where each of the four half circles is a heteroclinic X_γ-orbit connecting the two hyperbolic fixed points.

	hyperbolic fixed points	*critical value*
$\gamma < -1$	$(\pm r, 0, 0)$	$(1+\gamma)r^2$
$-1 < \gamma < 0$	$(0, 0, \pm r)$	0
$\gamma > 0$	$(0, \pm r, 0)$	γr^2

(5.4)

Let us verify the claimed classification of the two fixed points $(0, 0, \varepsilon r)$ with $\varepsilon \in \{\pm\}$. The other four fixed points are treated in a similar fashion. Near $(0, 0, \varepsilon r)$ we choose M, J as coordinates of $\mathbb{S}^2_r$. The equations of motion induced by X_γ on $\mathbb{S}^2_r$ in these coordinates read

$$\begin{aligned} \dot{M} &= -\varepsilon 2\gamma J \sqrt{r^2 - M^2 - J^2}, \\ \dot{J} &= \varepsilon 2(1+\gamma) M \sqrt{r^2 - M^2 - J^2}. \end{aligned}$$

If linearized at $(0, 0, \varepsilon r)$ the corresponding linear system is given by the 2×2-matrix εA where

$$A = 2r \begin{pmatrix} 0 & -\gamma \\ 1+\gamma & 0 \end{pmatrix}.$$

The eigenvalues of A are given by

$$\lambda^2 + 4\gamma(1+\gamma)r^2 = 0 \quad \text{or} \quad \lambda_{1,2} = \pm 2r\sqrt{-\gamma(1+\gamma)}.$$

As $-\gamma(1+\gamma) > 0$ iff $-1 < \gamma < 0$ it follows that $\lambda_{1,2}$ are in $\mathbb{R} \setminus \{0\}$ and hence that $(0, 0, \pm r)$ are both hyperbolic fixed points for $-1 < \gamma < 0$ whereas they are both elliptic if $\gamma < -1$ or $\gamma > 0$. For $-1 < \gamma < 0$, the inverse image $K_\gamma^{-1}(\{0\})$ is given by

$$\begin{aligned} K_\gamma^{-1}(\{0\}) &= \{(M,J,L) | (1+\gamma)M^2 + \gamma J^2 = 0; M^2 + J^2 + L^2 = r^2\} \\ &= \left\{ (M,J,L) | M = \pm\sqrt{\left|\frac{\gamma}{1+\gamma}\right|} J; M^2 + J^2 + L^2 = r^2 \right\}, \end{aligned}$$

whereas for $\gamma < -1$ or $\gamma > 0$, $K_\gamma^{-1}(\{0\}) = \{(0,0,\pm r)\}$.

5.3 Foliation of $\mathcal{P}_k$ for $0 < k < \frac{N}{4}$

In this section we present a detailed study of the geometry of the moment map $\mathcal{M} : T^*\mathbb{R}^4 \to \mathbb{R}^4$ defined by the integrable system $\mathcal{P}_k$ with commuting integrals $\mathcal{H}_k$, $\mathcal{H}_{\tilde{k}}$, G_k, K_k for $1 \le k < \frac{N}{4}$. As before we restrict to FPU chains with potential V whose expansion (1.4) satisfies $\beta \neq \alpha^2$. We show that in this case the vector field induced by K_k exhibits hyperbolic dynamics. It is convenient to introduce the following notation. Denote the standard coordinates of $T^*\mathbb{R}^4$ by (x,y) with $x = (x_i)_{1\le i\le 4}$ and $y = (y_i)_{1\le i\le 4}$, and introduce the action variables $I_j = \frac{1}{2}(x_j^2 + y_j^2)$ $(1 \le j \le 4)$, as well as the Hopf variables $(M_i, J_i, L_i)_{1\le i\le 2}$ given by

$$\begin{aligned} (M_1, J_1, L_1) &= \frac{1}{2}(x_1y_2 - x_2y_1, x_1x_2 + y_1y_2, I_1 - I_2), \\ (M_2, J_2, L_2) &= \frac{1}{2}(x_3y_4 - x_4y_3, x_3x_4 + y_3y_4, I_3 - I_4). \end{aligned}$$

By Lemma 3.2.3, the Poisson brackets between the variables $(M_i, J_i, L_i)_{1\le i\le 2}$ are given by

$$\{M_i, J_i\} = -L_i, \ \{J_i, L_i\} = -M_i, \ \{L_i, M_i\} = -J_i$$

whereas all other brackets vanish.

The moment map $\mathcal{M}$ then takes the form

$$\mathcal{M} : T^*\mathbb{R}^4 \to \mathbb{R}^4, \quad (x,y) \mapsto (H_1, H_2, G, K_\gamma)$$

where

$$H_1 = \frac{1}{2}(I_1 + I_2); \quad H_2 = \frac{1}{2}(I_3 + I_4); \quad G = L_1 - L_2$$

and where K_γ is a scalar multiple of K_k, given by

$$K_\gamma = \sum_{i=1}^{2} \frac{1}{2} d_{i,\gamma}(M_i^2 + J_i^2) + (M_1M_2 - J_1J_2)$$

with the coefficients $d_{1,\gamma}$, $d_{2,\gamma}$ defined by

$$d_{1,\gamma} = \frac{\gamma + s_k^2}{s_{2k}}, \quad d_{2,\gamma} = \frac{\gamma + c_k^2}{s_{2k}}$$

and as before, $s_k = \sin\frac{k\pi}{N}$, $c_k = \cos\frac{k\pi}{N}$. (The definition of the integral G above differs from the one given in (5.1) by the integral $H_1 - H_2$ as $I_1 - I_3 = L_1 - L_2 + H_1 - H_2$.)

First note that the origin $(0,0)$ in $T^*\mathbb{R}^4$ is the only critical point of $\mathcal{M}$ with rank $d_{x,y}\mathcal{M} = 0$. Moreover, $\mathcal{M}^{-1}\{(0,0)\} = \{(0,0)\}$. Next observe that when restricted to $T^*\mathbb{R}^2 \times \{0\}$, one has $G = \frac{1}{2}(I_1 - I_2)$ and $K_\gamma = d_{1,\gamma}(H_1^2 - L_1^2)$, hence they are functions of I_1, I_2 alone and $\mathcal{M}|_{T^*\mathbb{R}^2\times\{0\}}$ may be replaced by the map $(x,y) \mapsto (I_1, I_2, 0, 0)$. The geometry of the latter map is the one of two uncoupled harmonic oscillators. The subspace $\{0\} \times T^*\mathbb{R}^2$ is treated similarly. It remains to study the restriction of $\mathcal{M}$ to $T^*\mathbb{R}^4 \setminus ((T^*\mathbb{R}^2 \times \{0\}) \cup (\{0\} \times T^*\mathbb{R}^2))$. The Hamiltonian vector fields of H_1 and H_2 induce a torus action on $T^*\mathbb{R}^2$. The corresponding symplectic reduction is given by the product of two Hopf maps,

$$\mathcal{F} : \mathbb{S}^3_{\rho_1} \times \mathbb{S}^3_{\rho_2} \to \mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}, (x,y) \mapsto (M_i, J_i, L_i)_{1\le i\le 2}$$

where for $i = 1,2$, $\mathbb{S}^3_{\rho_i} = \{H_i = \rho_i^2/4\}$ is a sphere in $T^*\mathbb{R}^2$ and $r_i = \rho_i^2/4 = \sqrt{M_i^2 + J_i^2 + L_i^2}|_{\mathbb{S}^3_{\rho_i}} = H_i|_{\mathbb{S}^3_{\rho_i}}$. The fibers of $\mathcal{F}$ are 2-dimensional tori, obtained by the $\mathbb{S}^1 \times \mathbb{S}^1$-action of $H_1 \times H_2$. The reduced system is then given by $(\mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}, Y, X_\gamma)$, where Y and X_γ denote the reduced Hamiltonian vector fields induced by G and K_γ, respectively. To compute Y and X_γ, note that the equations of motion in the reduced system, corresponding to a Hamiltonian H, are given by

$$\frac{d}{dt}\begin{pmatrix} M_i \\ J_i \\ L_i \end{pmatrix} = \begin{pmatrix} M_i \\ J_i \\ L_i \end{pmatrix} \times \begin{pmatrix} \partial_{M_i} H \\ \partial_{J_i} H \\ \partial_{L_i} H \end{pmatrix}, \quad i = 1,2 \tag{5.5}$$

- see section 5.2 for details of this procedure. We then obtain

$$Y = \begin{pmatrix} J_1 \\ -M_1 \\ 0 \\ -J_2 \\ M_2 \\ 0 \end{pmatrix}, \quad X_\gamma = \begin{pmatrix} (J_2 - d_{1,\gamma}J_1)L_1 \\ (d_{1,\gamma}M_1 + M_2)L_1 \\ -(M_1J_2 + M_2J_1) \\ (J_1 - d_{2,\gamma}J_2)L_2 \\ (d_{2,\gamma}M_2 + M_1)L_2 \\ -(M_1J_2 + M_2J_1) \end{pmatrix}. \tag{5.6}$$

Further introduce the reduced moment map

$$\mathcal{M}_\gamma : \mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2} \to \mathbb{R}^2, (M_i, J_i, L_i)_{1\le i\le 2} \mapsto (G, K_\gamma).$$

We now study the critical points of $\mathcal{M}_\gamma$ with rank $d\mathcal{M}_\gamma = 0$, i.e. points of $\mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}$ which are fixed points of both Y and X_γ. From the expressions for Y and X_γ derived above, one easily sees that there are only four such critical points,

$$(M_i, J_i, L_i)_{1\le i\le 2} = \varepsilon(0,0,r_1,0,0,\pm r_2),$$

where $\varepsilon \in \{\pm\}$. The value of the critical point $\varepsilon(0,0,r_1,0,0,-r_2)$ by $\mathcal{M}_\gamma$ is $(\varepsilon(r_1 + r_2), 0)$, and

$$\mathcal{M}_\gamma^{-1}\{(\varepsilon(r_1 + r_2), 0)\} = \{\varepsilon(0,0,r_1,0,0,-r_2)\}.$$

Computing the Jacobian of the reduced vector field X_γ at the critical points one sees that they are elliptic fixed points of X_γ. The values of the other two critical points $\varepsilon(0,0,r_1,0,0,r_2)$ by $\mathcal{M}_\gamma$ are $(\varepsilon(r_1-r_2),0)$. The inverse image of $(\varepsilon(r_1-r_2),0)$ might have several connected components, depending on the values of γ and the additional bifurcation parameter

$$r := \frac{r_1}{r_2} > 0.$$

Our main results concerning the critical points $\varepsilon(0,0,r_1,0,0,r_2)$ are collected in the following theorem.

Theorem 5.3.1. *Assume that* $1 \le k < \frac{N}{4}$, $0 < r \le 1$, $\varepsilon \in \{\pm\}$, *and* $\gamma \in \mathbb{R}$. *The critical point* $\varepsilon(0,0,r_1,0,0,r_2)$ *of* $\mathcal{M}_\gamma$ *is a* hyperbolic *fixed point of the vector field* X_γ *if and only if*

$$\left|(\gamma + s_k^2)\sqrt{r} + (\gamma + c_k^2)\frac{1}{\sqrt{r}}\right| < 2 s_{2k}. \tag{5.7}$$

Otherwise it is an elliptic *fixed point of* X_γ. *If (5.7) is satisfied, the stable and unstable manifolds of* $\varepsilon(0,0,r_1,0,0,r_2)$ *both have dimension two. In the case* $r < 1$, *the connected component of* $\mathcal{M}_\gamma^{-1}\{\varepsilon(r_1-r_2,0)\}$ *containing* $\varepsilon(0,0,r_1,0,0,r_2)$ *is a 2-dimensional torus pinched at* $\varepsilon(0,0,r_1,0,0,r_2)$ *and consists of homoclinic* X_γ*-orbits. In the case* $r = 1$, $\mathcal{M}_\gamma^{-1}\{(0,0)\}$ *is a 2-dimensional torus pinched at the two points* $\pm(0,0,r_1,0,0,r_1)$, *and* $\mathcal{M}_\gamma^{-1}\{(0,0)\} \setminus \{\pm(0,0,r_1,0,0,r_1)\}$ *consists of heteroclinic* X_γ*-orbits.*

To prove Theorem 5.3.1 we separately treat for any given $1 \le k < \frac{N}{4}$ three subsets of the domain of the parameters γ and r. The results for these three cases are stated in detail in Propositions 5.3.2, 5.3.3, and 5.3.4 below.

Note that the inverse image $\mathcal{M}_\gamma^{-1}\{(\varepsilon(r_1-r_2),0)\}$ is invariant under the action of the vector field Y. The orbits of this action can be easily described,

$$\{(R_{-\phi}(M_1,J_1), L_1, R_\phi(M_2,J_2), L_2) \big| |\phi| \le \pi\}, \tag{5.8}$$

where $R_\phi(u,v)$ denotes the image of $(u,v) \in \mathbb{R}^2$ of the rotation R_ϕ by the angle ϕ in counterclockwise orientation. Hence given L_1 and L_2 with $|L_i| < r_i$ for $i = 1,2$ there exists a unique point $(\hat{M}_i, \hat{J}_i, L_i)_{1\le i\le 2}$ on such an orbit with the property that

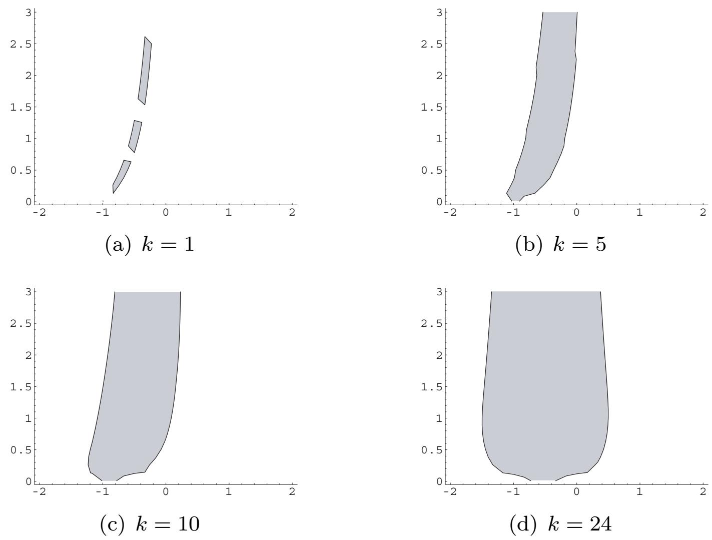
$$(\hat{M}_1, \hat{J}_1) = \left(\sqrt{r_1^2 - L_1^2}, 0\right) \quad \text{and} \quad (\hat{M}_2, \hat{J}_2) = \sqrt{r_2^2 - L_2^2}\ (\cos\alpha, \sin\alpha)$$

for some $0 \le \alpha < 2\pi$. We denote the corresponding Y-orbit by $\mathcal{L}(L_1, L_2, \alpha)$, i.e.

$$\mathcal{L}(L_1, L_2, \alpha) = \left\{\left(R_{-\phi}(\sqrt{r_1^2 - L_1^2}, 0), L_1, R_{\alpha+\phi}(\sqrt{r_2^2 - L_2^2}, 0), L_2\right) \big| |\phi| \le \pi\right\}. \tag{5.9}$$

As G and K_γ commute, K_γ is invariant along any orbit of the vector field Y, and we conclude that

$$K_\gamma((M_i, J_i, L_i)_{1\le i\le 2}) = \frac{1}{2}\sum_{i=1}^{2} d_{i,\gamma}(r_i^2 - L_i^2) + \sqrt{(r_1^2 - L_1^2)(r_2^2 - L_2^2)}\cos\alpha. \tag{5.10}$$



(a) $k=1$ (b) $k=5$

(c) $k=10$ (d) $k=24$

Figure 5.1: Subsets of parameters (γ, r) with hyperbolic dynamics of X_γ for $N=100$ and $k=1,5,10,24$

Let us now determine $\mathcal{M}_\gamma^{-1}\{(\varepsilon(r_1-r_2),0)\}$ for $\varepsilon=1$ and $r\leq 1$. (The case where $r>1$ and/or $\varepsilon=-1$ is treated in a similar fashion.) Let $(M_i,J_i,L_i)_{1\leq i\leq 2}$ be an element of $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}\setminus\{(0,0,r_1,0,0,r_2)\}$. Then $r_2-L_2=r_1-L_1$ and $K_\gamma((M_i,J_i,L_i)_{1\leq i\leq 2})=0$. First note that $L_1<r_1$. Indeed, if $L_1=r_1$, then $L_2=r_2$ and $(M_i,J_i)=(0,0)$ for $i=1,2$, contradicting our assumption on the point considered. Hence in the expression for K_γ displayed above we can factor out r_1-L_1 and the equation $K_\gamma=0$ reads

$$0=d_{1,\gamma}(r_1+L_1)+d_{2,\gamma}(r_2+L_2)+2\sqrt{(r_1+L_1)(r_2+L_2)}\cos\alpha. \tag{5.11}$$

Next let us consider the case where $L_1=-r_1$. Then $(M_1,J_1,L_1)=-(0,0,r_1)$ and (5.11) reads $(\gamma+c_k^2)(r_2+L_2)=0$. Hence either $L_2=-r_2$ or $\gamma=-c_k^2$. In the case $L_2=-r_2$ it follows from $r_2-L_2=r_1-L_1=2r_1$ that $r_2=r_1$ and $L_2=L_1$. As a consequence $(M_2,J_2,L_2)=-(0,0,r_1)$. On the other hand, if $\gamma=-c_k^2$ and $r_1<r_2$, then

$$-r_2<r_2-2r_1=L_2<r_2 \quad\text{and}\quad M_2^2+J_2^2=r_2^2-L_2^2=4r_1(r_2-r_1).$$

If $L_1\neq -r_1$ (and hence $L_2\neq -r_2$ as $r_1\leq r_2$) we set for $-r_1<L_1<r_1$

$$Q(L_1):=\begin{cases}\sqrt{\frac{r_1+L_1}{2r_2+L_1-r_1}} & \text{if } r_1<r_2\\ 1 & \text{if } r_1=r_2\end{cases} \tag{5.12}$$

Then $0<Q<\sqrt{r}$ and for $r<1$, Q is monotonically increasing on $-r_1<L_1<r_1$. After division by $\frac{1}{s_{2k}}\sqrt{(r_1+L_1)(r_2+L_2)}$ the equation (5.11) reads

$$(\gamma+s_k^2)Q+(\gamma+c_k^2)\frac{1}{Q}+2s_{2k}\cos\alpha=0. \tag{5.13}$$

To investigate the solutions of (5.13) we distinguish between three cases: $r = 1$, $[\gamma + c_k^2 = 0$ and $r < 1]$, and $[\gamma + c_k^2 \neq 0$ and $r < 1]$.

Let us first treat the case $r = 1$. Then $Q \equiv 1$ and equation (5.13) takes the form

$$2\gamma + 1 = -2s_{2k} \cos\alpha, \tag{5.14}$$

which is independent of L_1.

Proposition 5.3.2. *Let $\gamma \in \mathbb{R}$ be arbitrary and assume $1 \leq k < \frac{N}{4}$ and $r = 1$. Then the following statements hold:*

(i) If $|2\gamma + 1| > 2s_{2k}$, then

$$\mathcal{M}_\gamma^{-1}\{(0,0)\} = \{\varepsilon(0,0,r_1,0,0,r_1) | \varepsilon = \pm\},$$

and $\pm(0,0,r_1,0,0,r_1)$ are both elliptic *fixed points of the vector field X_γ.*

(ii) If $|2\gamma + 1| < 2s_{2k}$, then $\mathcal{M}_\gamma^{-1}\{(0,0)\} = \mathcal{N}_\alpha \cup \mathcal{N}_{-\alpha}$, where α is the unique angle satisfying $0 < \alpha < \pi$ and $2\gamma + 1 = -2s_{2k}\cos\alpha$, and where for any $-\pi \leq \beta \leq \pi$

$$\mathcal{N}_\beta = \bigcup_{\substack{|L_1| \leq r_1 \\ L_2 = L_1}} \mathcal{L}(L_1, L_2, \beta)$$

with $\mathcal{L}(L_1, L_2, \beta)$ given by (5.9). Both points, $\pm(0,0,r_1,0,0,r_1)$, are hyperbolic *fixed points of X_γ, and their stable and unstable manifolds have each dimension two. The set $\mathcal{N}_\alpha \setminus \{\pm(0,0,r_1,0,0,r_1)\}$ consists of heteroclinic X_γ-orbits from $(0,0,r_1,0,0,r_1)$ to $-(0,0,r_1,0,0,r_1)$, whereas $\mathcal{N}_{-\alpha} \setminus \{\pm(0,0,r_1,0,0,r_1)\}$ consists of heteroclinic X_γ-orbits with opposite direction. Topologically, $\mathcal{M}_\gamma^{-1}\{(0,0)\}$ is a 2-dimensional torus, pinched at each of the two fixed points $\pm(0,0,r_1,0,0,r_1)$.*

(iii) If $2\gamma + 1 = -2s_{2k}$, then $\alpha = 0$ and $\mathcal{M}_\gamma^{-1}\{(0,0)\} = \mathcal{N}_0$, whereas if $2\gamma + 1 = 2s_{2k}$, then $\alpha = \pi$ and $\mathcal{M}_\gamma^{-1}\{(0,0)\} = \mathcal{N}_\pi$. In both cases, $\pm(0,0,r_1,0,0,r_1)$ are elliptic *fixed points of X_γ. On $\mathcal{N}_0 \cup \mathcal{N}_\pi$, any X_γ-orbit is periodic and coincides with the corresponding Y-orbit at least up to orientation.*

Proof. (i) By a straightforward computation one shows that under the given assumptions, both points $\pm(0,0,r_1,0,0,r_1)$ are elliptic fixed points of X_γ. In view of equation (5.14), item (i) then easily follows. We give the details of this computation and all analogous computations in the proofs of Propositions 5.3.2-5.3.4 in Appendix F.

(ii) By the discussion preceding Proposition 5.3.2 it follows that the inverse image $\mathcal{M}_\gamma^{-1}\{(0,0)\}$ is given as claimed. Again by a straightforward computation one shows that both fixed points $\pm(0,0,r_1,0,0,r_1)$ are hyperbolic. To see that $\mathcal{N}_\alpha \setminus \{\pm(0,0,r_1,0,0,r_1)\}$ consists of heteroclinic orbits of the vector field X_γ, consider the third component $(X_\gamma)_3$ of X_γ (cf (5.6)). Any element $(M_i, J_i, L_i)_{1 \leq i \leq 2}$ in $\mathcal{N}_\alpha$ is of the form

$$(M_1, J_1) = \sqrt{r_1^2 - L_1^2}(\cos\phi, -\sin\phi),\ (M_2, J_2) = \sqrt{r_1^2 - L_1^2}(\cos(\alpha + \phi), \sin(\alpha + \phi)). \tag{5.15}$$

Thus

$$\begin{aligned}(X_\gamma)_3 &= -(M_1J_2 + M_2J_1) \\ &= -(r_1^2 - L_1^2)\,(\cos\phi\,\sin(\alpha+\phi) - \cos(\alpha+\phi)\sin\phi). \qquad (5.16)\end{aligned}$$

Hence $(X_\gamma)_3 = -(r_1^2 - L_1^2)\sin\alpha < 0$ for any point in $\mathcal{N}_\alpha \setminus \{\pm(0,0,r_1,0,0,r_1)\}$. As the last component of X_γ coincides with the third one, it follows that any X_γ-orbit on $\mathcal{N}_\alpha \setminus \{\pm(0,0,r_1,0,0,r_1)\}$ originates from $(0,0,r_1,0,0,r_1)$ and ends in $-(0,0,r_1,0,0,r_1)$. The orbits on $\mathcal{N}_{-\alpha}$ are analyzed in a similar way.

(iii) Clearly, if $2\gamma+1 = -2s_{2k}$, one has $\mathcal{M}_\gamma^{-1}\{(0,0)\} = \mathcal{N}_0$, and one verifies in a straightforward way that $\pm(0,0,r_1,0,0,r_1)$ are elliptic fixed points of X_γ. According to (5.16), the third component $(X_\gamma)_3$ of X_γ vanishes identically on $\mathcal{N}_0$. Further, $2\gamma+1 = -2s_{2k}$ implies that $1+d_{2,\gamma} = -1-d_{1,\gamma}$. In view of (5.15) it then follows that

$$X_\gamma = (1+d_{2,\gamma})L_2 \cdot Y.$$

The claimed statements for the case $2\gamma+1 = 2s_{2k}$ are proved in a similar fashion. □

Next we consider the case where $r < 1$ and $\gamma + c_k^2 = 0$. Then $\gamma + s_k^2 = -c_{2k}$ and hence

$$d_{1,\gamma} = -\frac{c_{2k}}{s_{2k}} \quad \text{and} \quad d_{2,\gamma} = 0.$$

Thus equation (5.13) takes the form

$$c_{2k}Q(L_1) = 2s_{2k}\cos\alpha. \qquad (5.17)$$

Note that $0 < c_{2k} < 1$ as $1 \le k < \frac{N}{4}$.

Proposition 5.3.3. *Assume that $1 \le k < \frac{N}{4}$, $0 < r < 1$, and $\gamma + c_k^2 = 0$. Then the following statements hold:*

(i) If $\sqrt{r} > 2s_{2k}/c_{2k}$, then the connected component of $\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\}$ containing the critical point $(0,0,r_1,0,0,r_2)$ consists of this point alone. It is an elliptic *fixed point of X_γ.*

(ii) If $\sqrt{r} \le 2s_{2k}/c_{2k}$, then

$$\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\} = \bigcup_{\substack{|L_1| \le r_1 \\ L_2 = L_1 + r_2 - r_1}} \mathcal{L}(L_1, L_2, \alpha_{L_1}) \cup \mathcal{L}(L_1, L_2, -\alpha_{L_1})$$

where for any $|L_1| \le r_1$, α_{L_1} is the unique angle satisfying

$$c_{2k}Q(L_1) = 2s_{2k}\cos\alpha_{L_1} \quad \text{and} \quad 0 \le \alpha_{L_1} \le \frac{\pi}{2}$$

and $Q(L_1)$ denotes the function defined by (5.12), continuously extended to the closed interval $[-L_1, L_1]$. Furthermore, the connected component of

$\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\}$ *containing* $(0,0,r_1,0,0,r_2)$ *consists of homoclinic* X_γ-*orbits which originate and end in* $(0,0,r_1,0,0,r_2)$. *Topologically, it is a* 2-*dimensional torus, pinched at* $(0,0,r_1,0,0,r_2)$.

If $\sqrt{r} < 2s_{2k}/c_{2k}$, *then* $(0,0,r_1,0,0,r_2)$ *is a* hyperbolic *fixed point of* X_γ *and its stable and unstable manifold have each dimension two. If* $\sqrt{r} = 2s_{2k}/c_{2k}$, $(0,0,r_1,0,0,r_2)$ *is an* elliptic *fixed point of* X_γ.

Proof. (i) By a straightforward computation one shows that under the given assumptions, $(0,0,r_1,0,0,r_2)$ is an elliptic fixed point of X_γ. In view of equation (5.17) and the discussion of the case $L_1 = \pm r_1$ item (i) then follows easily.

(ii) Again by the discussion preceding Proposition 5.3.2 it follows that the inverse image $\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\}$ is given as claimed. Again by a straightforward computation one sees that $(0,0,r_1,0,0,r_2)$ is a hyperbolic fixed point of X_γ if $\sqrt{r} < 2s_{2k}/c_{2k}$, and an elliptic one if $\sqrt{r} = 2s_{2k}/c_{2k}$. Next consider a point $(M_i, J_i, L_i)_{1\le i\le 2}$ in $\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\}$ with

$$\begin{aligned}(M_1, J_1) &= \sqrt{r_1^2 - L_1^2}\,(\cos\phi, -\sin\phi),\\ (M_2, J_2) &= \sqrt{r_2^2 - L_2^2}\,(\cos(\alpha_{L_1} + \phi), \sin(\alpha_{L_1} + \phi))\end{aligned}$$

where $|L_1| < r_1$. Then the third component of X_γ is given by (cf (5.16))

$$(X_\gamma)_3 = -\sqrt{r_1^2 - L_1^2}\,\sqrt{r_2^2 - L_2^2}\,\sin\alpha_{L_1}.$$

Hence $(X_\gamma)_3 = (X_\gamma)_6 < 0$. It follows that the X_γ-orbit passing through such a point originates at $(0,0,r_1,0,0,r_2)$ and then reaches a point of the form $(0,0,-r_1,M_2,J_2,L_2)$ with

$$L_2 = r_2 - 2r_1 > -r_2 \quad \text{and} \quad (M_2, J_2) = \sqrt{r_2^2 - L_2^2}(\cos(\pi + \tilde{\phi}), \sin(\pi + \tilde{\phi})). \tag{5.18}$$

At this point the vector field X_γ is given by

$$(-r_1 J_2, -r_1 M_2, 0, 0, 0, 0).$$

Note that this vector does not vanish as $M_2^2 + J_2^2 = r_2^2 - L_2^2 > 0$. Similarly, at a point $(M_i, J_i, L_i)_{1\le i\le 2}$ in $\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\}$ satisfying

$$\begin{aligned}(M_1, J_1) &= \sqrt{r_1^2 - L_1^2}\,(\cos\phi, -\sin\phi),\\ (M_2, J_2) &= \sqrt{r_2^2 - L_2^2}\,(\cos(-\alpha_{L_1} + \phi), \sin(-\alpha_{L_1} + \phi))\end{aligned}$$

and $|L_1| < r_1$ one has

$$(X_\gamma)_3 = \sqrt{r_1^2 - L_1^2}\,\sqrt{r_2^2 - L_2^2}\,\sin\alpha_{L_1}.$$

Hence $(X_\gamma)_3 = (X_\gamma)_6 > 0$. It follows that the X_γ-orbit passing through such a point ends up at $(0,0,r_1,0,0,r_2)$ and passes through a point of the form

$(0,0,-r_1,M_2,J_2,L_2)$ with (M_2,J_2,L_2) as in (5.18). We then conclude that the connected component of $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}$ containing $(0,0,r_1,0,0,r_2)$ consists of homoclinic X_γ-orbits originating and ending at $(0,0,r_1,0,0,r_2)$. □

Finally let us treat the case $r<1$ and $\gamma+c_k^2\neq 0$. Denote by $Q(L_1)$ the function defined by (5.12), extended continuously to the closed interval $[-r_1,r_1]$. Further introduce the function

$$f:(0,\sqrt{r})\to\mathbb{R},\ q\mapsto(\gamma+s_k^2)q+(\gamma+c_k^2)\frac{1}{q}. \tag{5.19}$$

Note that $\lim_{q\searrow 0}|f(q)|=\infty$.

Proposition 5.3.4. *Assume that* $1\leq k<\frac{N}{4}$, $0<r<1$, *and* $\gamma+c_k^2\neq 0$. *Then the following statements hold:*

(i) If $|f(\sqrt{r})|\geq 2s_{2k}$, *then the connected component of* $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}$ *containing the critical point* $(0,0,r_1,0,0,r_2)$ *consists of this point alone. It is an* elliptic *fixed point of* X_γ.

(ii) If $|f(\sqrt{r})|<2s_{2k}$, *then there exists* $-r_1<l_{\gamma,r}<r_1$ *so that the connected component of* $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}$ *containing* $(0,0,r_1,0,0,r_2)$ *is given by*

$$\bigcup_{\substack{l_{\gamma,r}\leq L_1\leq r_1\\ L_2=L_1+r_2-r_1}}\mathcal{L}(L_1,L_2,\alpha_{L_1})\cup\mathcal{L}(L_1,L_2,-\alpha_{L_1})$$

where for any $l_{\gamma,r}\leq L_1\leq r_1$, α_{L_1} *is the unique angle satisfying* $0\leq\alpha_{L_1}\leq\pi$ *and*

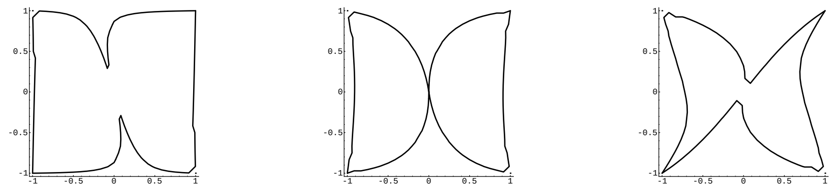
$$f(Q(L_1))=-2s_{2k}\cos(\alpha_{L_1}).$$

The point $(0,0,r_1,0,0,r_2)$ *is a* hyperbolic *fixed point of* X_γ *and its stable and unstable manifold each have dimension two. The connected component of* $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}$ *containing* $(0,0,r_1,0,0,r_2)$ *consists of homoclinic* X_γ-*orbits which originate and end in* $(0,0,r_1,0,0,r_2)$. *Topologically, it is a* 2-*dimensional torus, pinched at* $(0,0,r_1,0,0,r_2)$.

Proof. (i) By a straightforward computation one shows that under the given assumptions, $(0,0,r_1,0,0,r_2)$ is a (possibly degenerate) elliptic fixed point of X_γ. We have already seen that under the given assumption $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}\setminus\{(0,0,r_1,0,0,r_2)\}$ consists of the set of points $(M_i,J_i,L_i)_{1\leq i\leq 2}$ satisfying $L_2=L_1+r_2-r_1$ and (5.13). Note that equation (5.13) admits a solution α for $Q=\sqrt{r}$ iff $|f(\sqrt{r})|\leq s_{2k}$. In the case $|f(\sqrt{r})|>2s_{2k}$ it follows immediately that $\mathcal{M}_\gamma^{-1}\{(r_1-r_2,0)\}=\{(0,0,r_1,0,0,r_2)\}$. If $|f(\sqrt{r})|=2s_{2k}$, then an analysis of the graph of f near $(\sqrt{r},f(\sqrt{r}))$ leads to the claimed result.

(ii) As $\lim_{q\searrow 0}|f(q)|=\infty$ it follows that there exists $-r_1<l_{\gamma,r}<r_1$ so that the interval $[l_{\gamma,r},r_1]$ is a connected component of $(f\circ Q)^{-1}([-2s_{2k},2s_{2k}])$. It follows that for any $l_{\gamma,r}\leq L_1\leq r_1$ there exists a unique angle $0\leq\alpha_{L_1}\leq\pi$ so that

$$f(Q(L_1))=-2s_{2k}\cos(\alpha_{L_1}).$$



(a) $k = 1$ (b) $k = 3$ (c) $k = 5$

Figure 5.2: Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 1$, $\gamma = -1.35$

The connected component of the preimage $\mathcal{M}_\gamma^{-1}\{(r_1 - r_2, 0)\}$ containing the point $(0, 0, r_1, 0, 0, r_2)$ is then given as claimed. Again by a straightforward computation one sees that $(0, 0, r_1, 0, 0, r_2)$ is a hyperbolic fixed point of X_γ. One then can argue as in the proof of item (ii) of Proposition 5.3.3 to show the remaining claims. $\square$

Proof of Theorem 5.3.1. Theorem 5.3.1 follows from Propositions 5.3.2 - 5.3.4. $\square$

It remains to study the critical points of $\mathcal{M}_\gamma$ with rank $d\mathcal{M}_\gamma = 1$, i.e. points of $(\mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}) \setminus \{\pm(0, 0, r_1, 0, 0, \pm r_2)\}$ where the vector fields Y and X_γ are collinear. In view of the formulas (5.6) for Y and X_γ, points $(M_i, J_i, L_i) \in S^2_{r_i}$, $i = 1, 2$, of this type have the property that the determinant of any 2×2-submatrix of the 2×4-matrix formed by Y and X_γ vanishes. This leads to the following system of equations:

$$M_1 J_2 + M_2 J_1 = 0, \tag{5.20}$$

$$J_1^2 L_2 + L_1 J_2^2 - J_1 J_2 (d_{1,\gamma} L_1 + d_{2,\gamma} L_2) = 0, \tag{5.21}$$

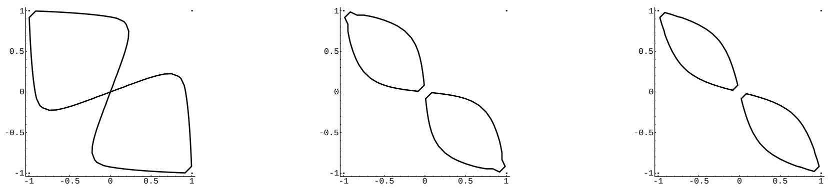
$$M_1^2 L_2 + L_1 M_2^2 + M_1 M_2 (d_{1,\gamma} L_1 + d_{2,\gamma} L_2) = 0, \tag{5.22}$$

$$M_1 J_1 L_2 - L_1 M_2 J_2 + J_1 M_2 (d_{1,\gamma} L_1 + d_{2,\gamma} L_2) = 0. \tag{5.23}$$

Theorem 5.3.5. *Assume that* $1 \le k < \frac{N}{4}$, $0 < r \le 1$, *and* $\gamma \in \mathbb{R}$. *If a point* $(M_i, J_i, L_i)_{1 \le i \le 2} \in \mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2} \setminus \{\pm(0, 0, r_1, 0, 0, \pm r_2)\}$ *is a critical point of* $\mathcal{M}_\gamma$ *with rank* $d\mathcal{M}_\gamma = 1$, *then* $(M_2, L_2) = \lambda(M_1, -J_1)$ *for some* $\lambda \in \mathbb{R}$, *and*

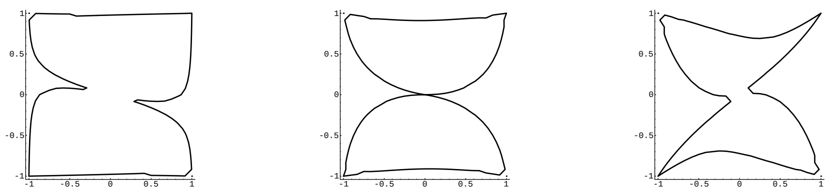
$$(r_1^2 - L_1^2)^2 L_2^2 + (r_2^2 - L_2^2)^2 L_1^2 + 2(r_1^2 - L_1^2)(r_2^2 - L_2^2)(2 L_1 L_2 - (d_{1,\gamma} L_1 + d_{2,\gamma} L_2)^2) = 0.$$

Given any point $(M_1, J_1, L_1) \in \mathbb{S}^2_{r_1} \setminus \{\pm(0, 0, r_1)\}$ *there exist at most eight points* $(M_2, J_2, L_2) \in \mathbb{S}^2_{r_2} \setminus \{\pm(0, 0, r_2)\}$ *such that* $(M_i, J_i, L_i)_{1 \le i \le 2}$ *is a critical point of* $\mathcal{M}_\gamma$ *with rank* $d\mathcal{M}_\gamma = 1$.



(a) $k = 1$ (b) $k = 3$ (c) $k = 5$

Figure 5.3: Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 1$, $\gamma = -0.5$



(a) $k = 1$ (b) $k = 3$ (c) $k = 5$

Figure 5.4: Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 1$, $\gamma = 0.35$

Proof. First assume that $L_1 \in \{\pm r_1\}$. Then $J_1 = M_1 = 0$. Hence (5.20) is automatically satisfied and equations (5.21) and (5.22) read $J_2 = 0$ and $M_2 = 0$, respectively. As a consequence, $(M_1, J_1, L_1) = (0, 0, \pm r_1)$ and $(M_2, J_2, L_2) = (0, 0, \pm r_2)$. In view of Theorem 5.3.1 we thus may assume that $|L_1| < r_1$. Then $(M_1, J_1) \neq (0, 0)$. Hence the first equation (5.20) says that there exists $\lambda \in \mathbb{R}$ such that

$$(M_2, J_2) = \lambda(M_1, -J_1). \tag{5.24}$$

The conditions $(M_i, J_i, L_i) \in \mathbb{S}^2_{r_i}$, $i = 1, 2$ then imply that λ satisfies

$$\lambda^2 = \frac{r_2^2 - L_2^2}{r_1^2 - L_1^2}. \tag{5.25}$$

Substituting (5.24) into (5.21)-(5.23) one sees, again using $(M_1, J_1) \neq (0, 0)$, that (5.21)-(5.23) is equivalent to

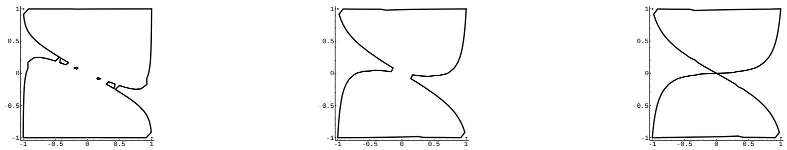
$$L_2 + \lambda^2 L_1 + \lambda(d_{1,\gamma} L_1 + d_{2,\gamma} L_2) = 0, \tag{5.26}$$

or, taking squares, $(L_2 + \lambda^2 L_1)^2 - \lambda^2(d_{1,\gamma} L_1 + d_{2,\gamma} L_2)^2 = 0$. Using (5.25), the latter equation reads

$$(r_1^2 - L_1^2)^2 L_2^2 + (r_2^2 - L_2^2)^2 L_1^2 + 2(r_1^2 - L_1^2)(r_2^2 - L_2^2)(2L_1L_2 - (d_{1,\gamma} L_1 + d_{2,\gamma} L_2)^2) = 0,$$

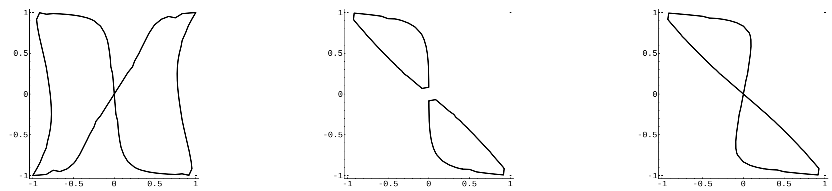
or, after dividing by $r_1^2 r_2^4$ one gets, using the bifurcation parameter r and the normed variables $l_i := L_i/r_i \in (0, 1)$ $(i = 1, 2)$,

$$r^2(1 - l_1^2)^2 l_2^2 + (1 - l_2^2)^2 l_1^2 + 2r(1 - l_1^2)(1 - l_2^2)(2l_1l_2 - (\sqrt{r} d_{1,\gamma} l_1 + \frac{1}{\sqrt{r}} d_{2,\gamma} l_2)^2) = 0. \tag{5.27}$$



(a) $k = 1$ (b) $k = 3$ (c) $k = 5$

Figure 5.5: Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 0.3$, $\gamma = -2.5$



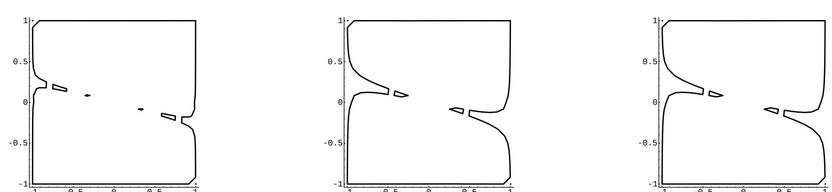
(a) $k = 1$ (b) $k = 3$ (c) $k = 5$

Figure 5.6: Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 0.3$, $\gamma = -1$

Note that for given r and $0 < l_1 < 1$, the left hand side of (5.27) is a polynomial in l_2 of degree four, i.e. $(l_1^2 + d_{2,\gamma}^2(1 - l_1^2))l_2^4 + O(l_2^3)$.

Summarizing, we have shown that for any given point in $\mathbb{S}^2_{r_1} \setminus \{(0, 0, \pm r_1)\}$, there exist at most eight points in $\mathbb{S}^2_{r_2} \setminus \{(0, 0, \pm r_2)\}$ such that Y and X_γ are collinear. Indeed for any $(M_1, J_1, L_1) \in \mathbb{S}^2_{r_1} \setminus \{(0, 0, \pm r_1)\}$, a solution $(M_2, J_2, L_2) \in \mathbb{S}^2_{r_2} \setminus \{(0, 0, \pm r_2)\}$ of (5.20)-(5.23) is given by $(M_2, J_2) = \lambda(M_1, J_1)$ and $L_2 = r_2 l_2$ with λ and l_2 satisfying (5.25) and (5.27), respectively. As a consequence the set of solutions of (5.20)-(5.23) is an algebraic subset of $\mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}$ of dimension at most two. $\square$

In order to analyze the critical points of $\mathcal{M}_\gamma$ with rank $d\mathcal{M}_\gamma = 1$, we perform



(a) $k = 1$ (b) $k = 3$ (c) $k = 5$

Figure 5.7: Sets of solutions (l_1, l_2) of (5.27) for $N = 24$, $r = 0.3$, $\gamma = 6$

another symplectic reduction. First we pass to the orbit space of the flow of Y on $\mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}$ and then to the level sets of G.

In view of (5.8), the Y-flow is an $\mathbb{S}^1$-action. Note that besides L_1 and L_2, the quantities σ and τ are invariant under this $\mathbb{S}^1$-action,

$$\sigma := M_1 M_2 - J_1 J_2, \quad \tau := M_1 J_2 + M_2 J_1.$$

They are related by the identity

$$\sigma^2 + \tau^2 = \prod_{i=1}^{2} (r_i^2 - L_i^2). \tag{5.28}$$

Define

$$\begin{array}{rccc} \mathcal{F}^{(3)}: & \dot{\mathbb{S}}^2_{r_1} \times \dot{\mathbb{S}}^2_{r_2} & \to & \mathbb{R}^4 \\ & (M_1, J_1, L_1, M_2, J_2, L_2) & \mapsto & (L_1, L_2, \sigma, \tau) \end{array}$$

where $\dot{\mathbb{S}}^2_{r_i} := \mathbb{S}^2_{r_i} \setminus \{(0, 0, \pm r_i)\}$. Let $\mathcal{O}_r$ denote the image of $\mathcal{F}^{(3)}$. For any element $(L_1, L_2, s, t) \in \mathcal{O}_r$ we have

$$s^2 + t^2 = \prod_{i=1}^{2} (r_i^2 - L_i^2) \quad \text{and} \quad |L_i| \le r_i \ (i = 1, 2). \tag{5.29}$$

The fibers of $\mathcal{F}^{(3)}$ are the orbits of the Y-action on $\dot{\mathbb{S}}^2_{r_1} \times \dot{\mathbb{S}}^2_{r_2}$, i.e. $\mathcal{O}_r$ coincides with the orbit space of the Y-action on $\dot{\mathbb{S}}^2_{r_1} \times \dot{\mathbb{S}}^2_{r_2}$. As a consequence, any function on $\dot{\mathbb{S}}^2_{r_1} \times \dot{\mathbb{S}}^2_{r_2}$ which Poisson commutes with G factors through $\mathcal{O}_r$.

In particular, K_γ and G factor through $\mathcal{O}_r$. In fact, K_γ and G, when expressed in the variables L_1, L_2, σ, τ, are polynomials, given by

$$K_\gamma = \sum_{i=1}^{2} \frac{1}{2} d_{i,\gamma} (r_i^2 - L_i^2) + \sigma, \tag{5.30}$$

$$G = L_1 - L_2. \tag{5.31}$$

By reducing the system (G, K_γ) by the Y-action one obtains a family of integrable systems with one degree of freedom parametrized by the value c of G. Denote by $X_{\gamma,c}$ the Hamiltonian vector field induced by K_γ. The fixed points of $X_{\gamma,c}$ can then be characterized in terms of the bifurcation parameters γ, r, and k.

Note that by (5.20), the rank-1-points of the reduced moment map $\mathcal{M}_\gamma$ satisfy $\tau = 0$, and by (5.24)-(5.25), $\sigma^2 = (r_2^2 - L_2^2)(r_1^2 - L_1^2)$. Hence the image of the set of the rank-1-points by $\mathcal{F}^{(3)}$ is an algebraic subset of $\mathcal{O}_r$ of dimension at most one - see (5.27).

By (5.28), σ and τ are located on a circle of radius $\sqrt{(r_1^2 - L_1^2)(r_2^2 - L_2^2)}$,

$$(\sigma, \tau) = \sqrt{(r_1^2 - L_1^2)(r_2^2 - L_2^2)}\, (\cos\phi, \sin\phi), \tag{5.32}$$

where $\phi \in \mathbb{R}/2\pi\mathbb{Z}$. The phase spaces, reduced by the Y-action, are now obtained by taking subsets of $\mathcal{O}_r$ corresponding to level sets of G, i.e. by replacing L_2 by $L_1 - c$, where c is the value of G. The restriction $K_{\gamma,c}$ of K_γ to the reduced phase space corresponding to the value c of G is then given by

$$\begin{aligned} K_{\gamma,c}(L_1,\phi) &= \frac{1}{2}\left(d_{1,\gamma}(r_1^2 - L_1^2) + d_{2,\gamma}(r_2^2 - (L_1-c)^2)\right) \\ &\quad + \sqrt{(r_1^2 - L_1^2)(r_2^2-(L_1-c)^2)}\,\cos\phi \end{aligned} \tag{5.33}$$

with $L_1 \in ((-r_1, r_1) \cap (c - r_2, c + r_2))$ and $\phi \in \mathbb{R}/2\pi\mathbb{Z}$.

The reduced Hamiltonian vector field induced by $K_{\gamma,c}$ is given by

$$X_{\gamma,c}(L_1,\phi) = \frac{d}{dt}\begin{pmatrix} L_1 \\ \phi \end{pmatrix} = \{L_1,\phi\}\begin{pmatrix} \partial K_{\gamma,c}/\partial\phi \\ -\partial K_{\gamma,c}/\partial L_1 \end{pmatrix}. \tag{5.34}$$

Note that

$$\begin{aligned} \{L_1,\phi\} &= \left\{L_1, \arctan\frac{\tau}{\sigma}\right\} = \frac{1}{1+(\tau/\sigma)^2}\left\{L_1,\frac{\tau}{\sigma}\right\} \\ &= \frac{1}{1+(\tau/\sigma)^2}\cdot\frac{\sigma\{L_1,\tau\} - \tau\{L_1,\sigma\}}{\sigma^2} = \frac{1}{1+(\tau/\sigma)^2}\cdot\frac{\sigma^2+\tau^2}{\sigma^2} = 1, \end{aligned}$$

since $\{L_1,\tau\} = \sigma$ and $\{L_1,\sigma\} = -\tau$. Furthermore, with $L_2 = L_1 - c$,

$$\begin{aligned} \frac{\partial K_{\gamma,c}}{\partial\phi} &= -\sqrt{(r_1^2-L_1^2)(r_2^2-L_2^2)}\,\sin\phi, \\ \frac{\partial K_{\gamma,c}}{\partial L_1} &= -(d_{1,\gamma}L_1 + d_{2,\gamma}L_2) - \frac{L_1(r_2^2-L_2^2)+L_2(r_1^2-L_1^2)}{\sqrt{(r_1^2-L_1^2)(r_2^2-L_2^2)}}\cos\phi. \end{aligned}$$

Hence (5.34) reads

$$X_{\gamma,c}(L_1,\phi) = \begin{pmatrix} -\sqrt{(r_1^2-L_1^2)(r_2^2-L_2^2)}\,\sin\phi \\ (d_{1,\gamma}L_1 + d_{2,\gamma}L_2) + \frac{L_1(r_2^2-L_2^2)+L_2(r_1^2-L_1^2)}{\sqrt{(r_1^2-L_1^2)(r_2^2-L_2^2)}}\cos\phi \end{pmatrix}, \tag{5.35}$$

where we treat $L_2 = L_1 - c$ as a dependent variable.

By (5.35), the fixed points of the vector field $X_{\gamma,c}$ with $|L_1| < r_1$ are given by (L_1,ϕ) satisfying

$$\phi \in \pi\mathbb{Z} \tag{5.36}$$

and

$$(d_{1,\gamma}L_1 + d_{2,\gamma}L_2) + \frac{L_1(r_2^2-L_2^2)+L_2(r_1^2-L_1^2)}{\sqrt{(r_1^2-L_1^2)(r_2^2-L_2^2)}}\cos\phi = 0. \tag{5.37}$$

Note that (5.36) and the square of (5.37) are equivalent to the system of equations (5.20)-(5.27) derived above.

In order to determine the type of the fixed points (L_1,ϕ), i.e. points satisfying (5.36)-(5.37) for a given value c of G, one computes the Jacobian of

$X_{\gamma,c}$, $H \equiv H_{\gamma,c}(L_1,\phi) = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$ at these points. Note that at such points $h_{11} = 0$ and $h_{22} = 0$ and thus $\det(H) = -h_{12}h_{21}$. Hence such a fixed point is an elliptic or hyperbolic fixed point of $X_{\gamma,c}$ if $h_{12}h_{21}$ is positive or negative. We investigate the signs of h_{12} and h_{21} separately.

First note that

$$h_{12} = -\sqrt{(r_1^2 - L_1^2)(r_2^2 - L_2^2)}\,\cos\phi,$$

hence, since $\phi \in \pi\mathbb{Z}$, $\operatorname{sign}(h_{12}) = -(-1)^{\phi/\pi}$. Next, we compute

$$h_{21} = \left(-(d_{1,\gamma} + d_{2,\gamma}) \pm \frac{\partial}{\partial L_1} \left(\frac{L_1(r_2^2 - (L_1 - c)^2) + L_2(r_1^2 - L_1^2)}{\sqrt{(r_1^2 - L_1^2)(r_2^2 - (L_1 - c)^2)}} \right) \right),$$

where the sign in front of the derivative is again equal to $-(-1)^{\phi/\pi}$. Another (lengthy) computation shows that the latter derivative is equal to the nonnegative expression

$$\frac{1}{\sqrt{(1-l_1^2)(1-l_2^2)}} \cdot \left(\sqrt{r \cdot \frac{1-l_1^2}{1-l_2^2}} - \left(\sqrt{r \cdot \frac{1-l_1^2}{1-l_2^2}} \right)^{-1} \right)^2,$$

again with $l_i := L_i/r_i$ $(i = 1,2)$. Since $d_{1,\gamma} + d_{2,\gamma} = \frac{2\gamma+1}{s_{2k}}$, the classification reduces investigating for solutions of (5.27) the sign of

$$\frac{2\gamma+1}{s_{2k}} \pm \frac{1}{\sqrt{(1-l_1^2)(1-l_2^2)}} \cdot \left(\sqrt{r \cdot \frac{1-l_1^2}{1-l_2^2}} - \left(\sqrt{r \cdot \frac{1-l_1^2}{1-l_2^2}} \right)^{-1} \right)^2.$$

Chapter 6

Discussion and Outlook

The fact that although at first glance, one-dimensional FPU chains appear to be a rather simple system, they exhibit at the same time such a rich dynamics, leads to several conclusions, some mathematical ones concerning perturbation theory in general and some physical or "general" ones concerning the FPU phenomena and their explanation. Let us first turn to the former issue.

First of all, it is the "simplicity" of the system under consideration which makes it possible to carry through all necessary computations in order to obtain normal forms of an order high enough to check the hypotheses of the KAM and Nekhoroshev theorems. Especially in the case of the latter theorem, it is precisely the difficulty of explicitly and rigorously checking its assumptions which seems to be the cause of the deplorable fact that the number of systems, to which the Nekhoroshev theorem has been applied, is rather small. In particular, it seems that up to now, there have not been many thorough discussions of Nekhoroshev's original criteria for "steepness" at a given example. Since these criteria apparently are considerably weaker than convexity or even quasi-convexity, such investigations could greatly extend the class of systems to which Nekhoroshev's estimates apply. Similar things can be said on Rüssmanns higher order nondegeneracy conditions, which are a weaker version of Kolmogorov's original nondegeneracy conditions (these weaker conditions have apparently been thought of in order to deal with systems which are obviously not nondegenerate).

Even though FPU chains fail to meet the original nondegeneracy conditions only in some exceptional cases of the parameter values, we think that it would be worthwhile to try to check these higher order conditions in these exceptional cases, and it seems possible that this could be accomplished by simply pursuing further the approach of this thesis, namely explicitly computing the coefficients of the Birkhoff normal form up to higher and higher orders and then checking the appropriate conditions. Of course, it cannot be taken for granted that there are no resonances, i.e. obstructions to the transformation to Birkhoff normal forms up to higher orders, similarly to the case of even periodic chains, where there are fourth order resonances leading to a *resonant* normal form of order four. However, due to the fact that we know that the (full) periodic Toda lattice is an integrable system, we are confident that it should be possible to carry thorugh this procedure at least in some cases of the parameter values, namely those approximating the Toda lattice up to higher and higher orders.

Besides justifying the claims that the KAM theorem can be applied to FPU chains in the cases of odd periodic and Dirichlet chains, we also have thoroughly investigated even periodic chains. First of all, it is a surprising fact that the truncated fourth order Hamiltonian of these chains turns out to be integrable for all parameter values. Even though we do not explain this integrability by abstract geometric or group-theoretic arguments, we think that it could also contribute to an explanation of the FPU phenomena, depending on the results of the planned numerical implementation of our results. Moreover, the detailed analysis of the level sets of the associated moment map reveals an extremely rich geometric structure, which is also somewhat surprising in view of the fact that we have partitioned all the integrals of this integrable system into subsets of at most four integrals, i.e. the systems under investigation "live" in an a phase space of dimension at most eight. Nevertheless, only after repeated reductions we have been able to properly classify the various critical points of the originally given moment map. Moreover, the bifurcation diagrams obtained at two different steps of this reduction process turn out to have a rich geometric structure themselves - we have tried to convey an impression of this structure by plotting some particularly interesting examples. However, there are a lot of questions concerning these bifurcations which we have not answered, for instance for which parameter values the "domains of hyperbolicity" are connected, what their asymptotic behavior is in the case of the particle number tending to infinity, just to mention some of these questions. Similar questions could be posed for the set of critical points of the moment map of rank one - not only concerning their nature (hyperbolic or elliptic), but also their distribution in the plane of the two (normed) action variables, what the geometric properties of the set given by the solutions of the appropriate equations are. And again, we emphasize that all these questions arise from the analysis of a system in an eight-dimensional phase space. It seems quite likely that similar systems on phase spaces of higher dimensions can become analytically intractable quite rapidly.

Returning to those FPU chains where we are able to compute Birkhoff normal forms of order four, the further computation of coefficients of an even higher order would also contribute to a preciser implementation of our transformation formulas - and this directly leads to the second issue to be discussed, namely the relevance of our results for the explanation of the FPU phenomena. As already emphasized in the introduction, before having implemented our transformation formulas, we do not attempt to fully answer the question of the "explanatory" power of our work, we have just rigorously justified the claim that odd periodic and in particular Dirichlet chains (as originally considered by Fermi, Pasta, and Ulam) can be considered as fifth-order perturbations of a nondegenerate integrable system, thereby confirming a conjecture which has been proposed repeatedly in the last 40 years. Furthermore and more generally, it seems to be quite promising that it has turned out to be possible to approximate all three types of chains (including the even periodic ones) with integrable systems up to fourth order. But, again, it is explicitly not our attempt to work "against" the other approach towards a resolution of the FPU paradox, namely investigating the continuous limit of the discrete chain and explaining the behavior of the discrete chain by the "soliton-like" behavior of the continuous limit.

In this way, we arrive at another promising direction for future research, namely to tackle the problems arising at the "interface" of the discrete and the continuous models. A first step could be to analyze how our results behave in the limit of the number of particles tending to infinity. Of course, we do not claim that this has to be started "ab ovo" - already many steps have been undertaken in this direction.

Finally, the FPU problem is a "toy example" insofar as it has been constructed from idealized assumptions, let us only mention the assumptions of equal masses and only nearest-neighbor interaction. Whereas this makes all our very explicit investigations possible, it has the drawback that one should be extremely careful in drawing "realistic" conclusions from our results - recall also the remarks in the introduction on the epistemolgical issues of the FPU problem. Nevertheless, we think that rigorously understanding such a comparatively simple system can be of great help towards the explanation of more complex systems (also, and in particular, non-physical systems).

Appendix A

Details of section 3.1

A.1 Relative coordinates

We begin by expressing the FPU Hamiltonian H_V in relative coordinates. Introduce $(v = (v_j)_{1\le j\le N-1}, v_N) \in \mathbb{R}^N$ given by (3.1). Then $(v, v_N) = Mq$ is the linear change of the coordinates $q_1, \ldots, q_N$ where M is given by

$$M = \begin{pmatrix} -1 & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & & & 0 \\ 0 & \ldots & 0 & -1 & 1 \\ N^{-1} & \ldots & & \ldots & N^{-1} \end{pmatrix}.$$

The variables $(u = (u_j)_{1\le j\le N-1}, u_N) \in \mathbb{R}^N$ conjugate to (v, v_N) are then given by $(M^T)^{-1}p$. The inverse of M^T, $(M^T)^{-1}$, can be computed to be

$$(M^T)^{-1} = \frac{1}{N}\begin{pmatrix} 1 & \ldots & \ldots & 1 \\ 2 & \ldots & \ldots & 2 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ N & \ldots & \ldots & N \end{pmatrix} - \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & & & & \vdots \\ 1 & \ldots & & 1 & 0 \\ 0 & \ldots & & \ldots & 0 \end{pmatrix}. \tag{A.1}$$

Note that by (A.1), $u_k = kP - \sum_{j=1}^k p_j$ for any $1 \le k \le N-1$ and $u_N = NP$. Hence

$$p_1 = -u_1 + P; \quad p_N = u_{N-1} + P; \quad p_k = (u_{k-1} - u_k) + P \qquad (2 \le k \le N-1)$$

and thus

$$\frac{1}{2}\sum_{j=1}^{N} p_j^2 = \frac{NP^2}{2} + \frac{1}{2}\left(u_1^2 + (u_1 - u_2)^2 + \ldots + (u_{N-2} - u_{N-1})^2 + u_{N-1}^2\right).$$

Moreover, using that $q_{N+1} - q_N = q_1 - q_N = -\sum_{k=1}^{N-1}(q_{k+1} - q_k)$ one gets for any $s \in \mathbb{Z}_{\geq 1}$

$$\sum_{j=1}^{N}(q_{j+1} - q_j)^s = \sum_{k=1}^{N-1} v_k^s + (-1)^s \left(\sum_{k=1}^{N-1} v_k\right)^s .$$

Combining the two expressions displayed above yields $H_V = \frac{NP^2}{2} + \tilde{H}_V$, where $\tilde{H}_V$ only depends on (v, u) and is given by

$$\tilde{H}_V = \frac{1}{2}\left(u_1^2 + \sum_{l=1}^{N-2}(u_{l+1} - u_l)^2 + u_{N-1}^2\right) + \frac{1}{2}\left(\sum_{k=1}^{N-1} v_k^2 + \left(\sum_{k=1}^{N-1} v_k\right)^2\right) \tag{A.2}$$
$$+\frac{\alpha}{3!}\left(\sum_{k=1}^{N-1} v_k^3 - \left(\sum_{k=1}^{N-1} v_k\right)^3\right) + \frac{\beta}{4!}\left(\sum_{k=1}^{N-1} v_k^4 + \left(\sum_{k=1}^{N-1} v_k\right)^4\right) + O(v^5).$$

Note that for any values of α and β, the point $(v, u) = (0, 0)$ is a critical point of the Hamiltonian $\tilde{H}_V$.

A.2 Birkhoff normal fom up to order two

To compute the Birkhoff normal form of $\tilde{H}_V$ up to order two near the fixed point $(v, u) = (0, 0)$, we take the expansion (A.2) as a starting point and substitute the transformation formulas (3.5)-(3.8). The following lemma gives a self-contained proof of the fact that this linear map is canonical.

Lemma A.1. *The linear transformation $\mathcal{Z} \to \mathbb{R}^{2N-2}$, $\zeta \mapsto (v, u)$, as defined by (3.5)-(3.8), is a canonical isomorphism.*

Proof. First let us show

$$\{v_l(\zeta), u_m(\zeta)\} = i\,\delta_{lm}, \tag{A.3}$$
$$\{v_l(\zeta), v_m(\zeta) = 0, \tag{A.4}$$
$$\{u_l(\zeta), u_m(\zeta)\} = 0 \tag{A.5}$$

for any $1 \leq l, m \leq N - 1$. Since (v, u) are canonical coordinates on $\mathbb{R}^{2N-2}$, the proof of (A.3) amounts to showing that

$$\sum_{k=1}^{N-1}\left(\frac{\partial v_l}{\partial \zeta_k}\frac{\partial u_m}{\partial \zeta_{-k}} - \frac{\partial v_l}{\partial \zeta_{-k}}\frac{\partial u_m}{\partial \zeta_k}\right) = i\,\delta_{lm}$$

for any $1 \leq l, m \leq N-1$. It follows from (3.5)-(3.8) that for any $1 \leq k \leq N-1$,

$$\frac{\partial v_l}{\partial \zeta_k} = \frac{\lambda_k}{\sqrt{N}} e^{\pi i(2l-1)k/N}, \quad \frac{\partial v_l}{\partial \zeta_{-k}} = \frac{\lambda_k}{\sqrt{N}} e^{-\pi i(2l-1)k/N},$$

$$\frac{\partial u_m}{\partial \zeta_k} = \frac{\lambda_k}{\sqrt{N}} \sum_{j=0}^{m-1} e^{2\pi ijk/N}, \quad \frac{\partial u_m}{\partial \zeta_{-k}} = \frac{\lambda_k}{\sqrt{N}} \sum_{j=0}^{m-1} e^{-2\pi ijk/N}.$$

Hence

$$\begin{aligned}
\frac{\partial v_l}{\partial \zeta_k}\frac{\partial u_m}{\partial \zeta_{-k}} &- \frac{\partial v_l}{\partial \zeta_{-k}}\frac{\partial u_m}{\partial \zeta_k} \\
&= \frac{\lambda_k^2}{N}\left(e^{\pi i(2l-1)k/N} \sum_{j=0}^{m-1} e^{-2\pi ijk/N} - e^{-\pi i(2l-1)k/N} \sum_{j=0}^{m-1} e^{2\pi ijk/N} \right) \\
&= \frac{1}{N} \sin\frac{k\pi}{N} \sum_{j=0}^{m-1} \left(e^{\frac{\pi ik}{N}(2l-2j-1)} - e^{\frac{\pi ik}{N}(2j-2l+1)} \right) \\
&= \frac{2i}{N} \sin\frac{k\pi}{N} \sum_{j=0}^{m-1} \sin\left(\frac{k\pi}{N}(2(l-j)-1) \right) \\
&= \frac{i}{N} \sum_{j=0}^{m-1} \left(\cos\frac{2k\pi(1-(l-j))}{N} - \cos\frac{2k\pi(l-j)}{N} \right),
\end{aligned}$$

where for the latter identity we used that $2\sin x \sin y = \cos(x-y) - \cos(x+y)$. Taking the sum over k and changing the order of summation then leads to

$$\begin{aligned}
\sum_{k=1}^{N-1}\left(\frac{\partial v_l}{\partial \zeta_k}\frac{\partial u_m}{\partial \zeta_{-k}} - \frac{\partial v_l}{\partial \zeta_{-k}}\frac{\partial u_m}{\partial \zeta_k} \right) &= \frac{i}{N} \sum_{j=0}^{m-1}\sum_{k=1}^{N-1} \left(\cos\frac{2k\pi(1-(l-j))}{N} - \cos\frac{2k\pi(l-j)}{N} \right) \\
&= \frac{i}{N} \sum_{j=0}^{m-1} N(\delta_{l-j,1} - \delta_{l-j,0}) \\
&= i \sum_{j=0}^{m-1} (\delta_{l,j+1} - \delta_{l,j}) = i(\delta_{lm} - \delta_{l0}) = i\delta_{lm},
\end{aligned}$$

as claimed. To prove (A.4) and (A.5) one argues in a similar way. From (A.3)-(A.5) it immediately follows that the linear map $\xi \mapsto (v, u)$ is a canonical isomorphism. □

We now compute $\tilde{H}_V$ in terms of the new variables ζ. Write $\tilde{H}_V$ as $\tilde{H}_V = H_u + H_v$ where H_u and H_v denote the u- and v-dependent parts of (A.2), respectively. We compute $H_u(\zeta)$ and $H_v(\zeta)$ separately. To obtain $H_u(\zeta)$, we substitute (3.5)-(3.7) into the expression $\frac{1}{2}\left(u_1^2 + \sum_{l=1}^{N-2}(u_{l+1} - u_l)^2 + u_{N-1}^2\right)$ and get

$$H_u(\zeta) = \frac{1}{2N} \sum_{l=0}^{N-1} \left(\sum_{1 \le |k| \le N-1} \lambda_k e^{2\pi ilk/N} \zeta_k \right)^2$$

$$= \frac{1}{2N} \sum_{1\le|k|,|k'|\le N-1} \lambda_k \lambda_{k'} \left(\sum_{l=0}^{N-1} e^{2\pi i l(k+k')/N} \right) \zeta_k \zeta_{k'}.$$

Using again that $\sum_{l=0}^{N-1} e^{2\pi i l k/N} = N\delta_{k0}$ and $\lambda_k = \lambda_{-k}$ for any $1 \le |k| \le N-1$, one obtains

$$H_u(\zeta) = \sum_{k=1}^{N-1} \sin\frac{k\pi}{N}\, \zeta_k \zeta_{-k}.$$

Before computing $H_v(\zeta)$, we simplify its expansion in terms of the variables $(v_k)_{1\le k\le N-1}$. Define v_0 by the expression on the right hand side of (3.8) evaluated at $l = 0$. Note that

$$\sum_{l=0}^{N-1} v_l = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \lambda_k \zeta_k e^{-i\pi k/N} \left(\sum_{l=0}^{N-1} e^{2\pi i l k/N} \right) = 0.$$

Hence $\sum_{l=1}^{N-1} v_l = -v_0$ and therefore

$$H_v = \sum_{l=0}^{N-1} \left(\frac{1}{2} v_l^2 + \frac{\alpha}{3!} v_l^3 + \frac{\beta}{4!} v_l^4 \right) + O(|v|^5). \tag{A.6}$$

Substituting the expression (3.8) for v_l in the quadratic term in the expansion (A.6), we get

$$\begin{aligned} \frac{1}{2}\sum_{l=0}^{N-1} v_l^2 &= \frac{1}{2N} \sum_{1\le|k|,|k'|\le N-1} \lambda_k \lambda_{k'} \left(\sum_{l=0}^{N-1} e^{2\pi i l(k+k')/N} \right) e^{-i\pi(k+k')/N} \zeta_k \zeta_{k'} \\ &= \sum_{k=1}^{N-1} \sin\frac{k\pi}{N}\, \zeta_k \zeta_{-k}, \end{aligned}$$

where we again used that $\lambda_k = \lambda_{-k}$ and $\sum_{l=0}^{N-1} e^{2\pi i l k/N} = N\delta_{k0}$ for any $0 \le |k| \le N-1$.

The terms of third and fourth order in H_v are treated similarly. Combining the above computations leads to the claimed formula

$$\tilde{H}_V(\zeta) = H_u(\zeta) + H_v(\zeta) = G_2 + \alpha G_3 + \beta G_4 + O(\zeta^5)$$

with G_2, G_3, and G_4 given by (3.9), (3.10), and (3.11), respectively.

Appendix B

Nonresonance lemma

For the convenience of the reader, we provide a detailed proof of Lemma 3.1.3 in this appendix. This lemma and its proof are due to Beukers and Rink - see ([70], Appendix A). A very similar statement, from which Lemma 3.1.3 can be deduced as well, has been proven by Conway and Jones - see [14]. Recall from (3.13) that $K_4 \setminus K_4^N \subseteq \mathbb{Z}^4$ denotes the subset of quadruples (k_1, k_2, k_3, k_4) satisfying $1 \leq |k_i| \leq N-1$ $(1 \leq i \leq 4)$ and $k_1 + k_2 + k_3 + k_4 \equiv 0 \bmod N$ so that there are no integers l, m with $\{l, m, -l, -m\} = \{k_1, k_2, k_3, k_4\}$, and

$$K_4^{res} := K_{res}^{+} \cup K_{res}^{-} \subseteq K_4$$

where

$$K_{res}^{\pm} := \Big\{(k_1, k_2, k_3, k_4) \in K_4 | \ \exists\, l \in \mathbb{N} : 1 \leq l \leq \frac{N}{4} \ \text{ so that}$$
$$\{k_1, k_2, k_3, k_4\} = \{\pm l, \pm l \mp N, \frac{N}{2} \mp l, -\frac{N}{2} \mp l\}\Big\}.$$

Note that $K_4^{res} = \emptyset$ if N is odd. Let us restate Lemma 3.1.3 as follows:

Lemma B.1. *Let $(k_1, k_2, k_3, k_4) \in K_4 \setminus K_4^N$. Then $(k_1, k_2, k_3, k_4) \in K_4^{res}$ if and only if*

$$\sin \frac{k_1 \pi}{N} + \sin \frac{k_2 \pi}{N} + \sin \frac{k_3 \pi}{N} + \sin \frac{k_4 \pi}{N} = 0.$$

Let us make a few preparations for the proof of Lemma B.1. By a straightforward computation one sees that the "only if"-part of the claimed equivalence holds:

Lemma B.2. *For any $(k_1, k_2, k_3, k_4) \in K_4^{res}$, one has $\sum_{i=1}^{4} \sin \frac{k_i \pi}{N} = 0$.*

So it remains to prove the converse. First we consider some special cases.

Lemma B.3. *Let $(k_1, k_2, k_3, k_4) \in K_4 \setminus (K_4^N \cup K_4^{res})$. If there exist $l, m, n \in \mathbb{Z}$ such that* *(i)* $\{k_1, k_2, k_3, k_4\} = \{l, -l, m, n\}$, *or*

(ii) $\{k_1, k_2, k_3, k_4\} = \{l, N-l, m, n\}$ *with* $1 \le l \le N-1$, *or*

(iii) $\{k_1, k_2, k_3, k_4\} = \{l, -N-l, m, n\}$ *with* $-(N-1) \le l \le -1$,

then

$$\sum_{i=1}^{4} \sin \frac{k_i \pi}{N} \neq 0.$$

Proof. In case (i), it follows that $m+n = N$ (and thus $1 \le m, n \le N-1$) or $m+n = -N$ (and thus $-(N-1) \le m, n \le -1$). Hence in both cases, $\sin \frac{m\pi}{N}$ and $\sin \frac{n\pi}{N}$ have the same sign and $\sum_{i=1}^4 \sin \frac{k_i\pi}{N} = \sin \frac{m\pi}{N} + \sin \frac{n\pi}{N} \neq 0$. In the case (ii), by assumption, $m+n \equiv 0 \mod N$. The case $m+n=0$ has already been treated under (i). If $m+n = N$, then $\sin \frac{k_i\pi}{N} > 0$ for any $1 \le i \le 4$. If $m+n = -N$, then $m < 0$, and $m \notin \{-l, -N+l\}$. Thus $n = -N-m < 0$ and therefore $\sum_{i=1}^4 \sin \frac{k_i\pi}{N} = 2\sin \frac{l\pi}{N} - 2\sin \frac{(-m)\pi}{N} \neq 0$. The case (iii) is treated similarly as (ii). □

Another special case is treated in the following lemma.

Lemma B.4. *Assume that* $(k_1, k_2, k_3, k_4) \in K_4 \setminus K_4^N$ *satisfies*

$$k_i + k_j \not\equiv 0 \ mod\ N \quad \forall\, 1 \le i, j \le 4. \tag{B.1}$$

If there exist $l, n \in \{k_1, k_2, k_3, k_4\}$ *with*

$$\sin \frac{l\pi}{N} + \sin \frac{n\pi}{N} = 0, \tag{B.2}$$

then

$$\sum_{i=1}^{4} \sin \frac{k_i \pi}{N} = 0 \tag{B.3}$$

implies that $(k_1, k_2, k_3, k_4) \in K_4^{res}$.

Proof. From the assumptions (B.1)-(B.2) it follows that there exists $1 \le l \le N-1$ so that $\{k_1, k_2, k_3, k_4\} = \{l, -N+l, m, n\}$ for some $m, n \in \mathbb{Z}$. Then $\sin \frac{l\pi}{N} + \sin \frac{(-N+l)\pi}{N} = 0$ and hence by (B.3), $\sin \frac{m\pi}{N} + \sin \frac{n\pi}{N} = 0$. W.l.o.g. assume that $1 \le m \le N-1$. Then either $n = -m$ or $n = -N+m$. If $n = -m$, then $(k_1, k_2, k_3, k_4) \in K_4^{res}$ by Lemma B.3 (i). If $n = -N+m$, then one has

$$\sum_{i=1}^{4} k_i = 2l - N + 2m - N = 2(l+m) - 2N.$$

Note that $2(l+m) - 2N$ cannot be an even multiple of N, as otherwise $l+m \equiv 0 \mod N$, violating (B.1). If, in addition, N is odd, then $2(l+m) - 2N$ cannot be an odd multiple of N. Hence in the case where N is odd we conclude that $\sum_{i=1}^4 k_i \not\equiv 0 \mod N$, contradicting the assumption $(k_1, k_2, k_3, k_4) \in K_4$.

If N is even, it is however possible that $2(l+m)-2N$ equals $\pm N$: If $2(l+m)-2N=N$, i.e. $l+m=\frac{3}{2}N$, it follows that $\frac{N}{2}<l,m\leq N-1$, and $(k_1,k_2,k_3,k_4)\in K^-_{res}$ as $\{k_1,k_2,k_3,k_4\}=\{-l',-l'+N,\frac{N}{2}+l',-\frac{N}{2}+l'\}$ with $l'=l-\frac{N}{2}$. If $2(l+m)-2N=-N$, i.e. $l+m=\frac{N}{2}$, it follows similarly that $(k_1,k_2,k_3,k_4)\in K^+_{res}$ as $\{k_1,k_2,k_3,k_4\}=\{l,l-N,\frac{N}{2}-l,-\frac{N}{2}-l\}$. So in both cases, we conclude that $(k_1,k_2,k_3,k_4)\in K_4^{res}$. □

In view of Lemma B.3 and Lemma B.4 in order to prove Lemma B.1 it remains to show the following

Lemma B.5. *Assume that* $(k_1,k_2,k_3,k_4)\in K_4$ *satisfies (B.1). If for any* $1\leq i,j\leq 4$

$$\sin\frac{k_i\pi}{N}+\sin\frac{k_j\pi}{N}\neq 0. \tag{B.4}$$

(and thus $(k_1,k_2,k_3,k_4)\notin K_4^N\cup K_4^{res}$*), then*

$$\sum_{i=1}^{4}\sin\frac{k_i\pi}{N}\neq 0.$$

To prove Lemma B.5 let us first rewrite (B.3), using Euler's formula for the sine function,

$$\sum_{1\leq|j|\leq 4}\zeta_j=0 \tag{B.5}$$

where $\zeta_{\pm j}=\pm e^{\pm ik_j\pi/N}$ are $2N$'th roots of unity. Note that for any quadruple $(k_1,k_2,k_3,k_4)\in K_4\setminus K_4^N$ satisfying (B.4) one has

$$\zeta_j+\zeta_{j'}\neq 0 \quad \forall\, 1\leq|j|\leq|j'|\leq 4.$$

Indeed for any $1\leq|j|\leq|j'|\leq 4$ one has Im ζ_j+ Im $\zeta_{j'}=\sin\frac{k_{|j|}\pi}{N}+\sin\frac{k_{|j'|}\pi}{N}$ which does not vanish by assumption (B.4).

Let us first discuss equation (B.5) and its solutions in general, i.e. we consider the equation

$$\zeta_1+\ldots+\zeta_8=0 \tag{B.6}$$

and want to study its solutions, $(\zeta_l)_{1\leq l\leq 8}$, on the unit circle $S^1:=\{z\in\mathbb{C}\big||z|=1\}$.

We need an auxiliary result which we discuss first. Let $n\geq 2$ be arbitrary and assume that the sequence $(\zeta_i)_{1\leq i\leq n}\subseteq S^1$ has no vanishing subsums (i.e. $\sum_{l\in J}\zeta_l\neq 0$ for any nonempty proper subset J of $\{1,\ldots,n\}$) and satisfies the equation

$$\sum_{i=1}^{n}\zeta_i=0. \tag{B.7}$$

Let $M\in\mathbb{N}$ be the smallest positive integer with the property that $(\zeta_i/\zeta_j)^M=1$ for all $1\leq i,j\leq n$. Then there exists $\xi\in S^1$ so that $\zeta_i^M=\xi^M$ for any

$1 \leq i \leq n$. W.l.o.g. we can assume that $\xi = 1$. Furthermore, let p^k be a prime power dividing M so that M/p^k and p are relatively prime and define

$$M' =: M/p \quad \text{and} \quad \eta := e^{2\pi i/p^k}. \tag{B.8}$$

Then for any $1 \leq l \leq n$ there exists a unique integer $\mu(l)$ with $0 \leq \mu(l) \leq p-1$ such that $\zeta_l = \tilde{\zeta}_l \cdot \eta^{\mu(l)}$ where $\tilde{\zeta}_l$ is an element of the field $K := \mathbb{Q}(e^{2\pi i/M'})$. (As $\zeta_l^M = 1$ there exists $0 \leq r_l \leq M-1$ with $\zeta_l = e^{\frac{2\pi i}{M} r_l}$. If $r_l \equiv 0 \bmod p$ choose $\mu(l) = 0$. If $r_l \not\equiv 0 \bmod p$ choose $1 \leq \mu(l) \leq p-1$ so that $r_l \equiv \frac{M}{p^k}\mu(l) \bmod p$.) Hence (B.7) can be written as

$$0 = \sum_{l=1}^{n} \zeta_l = \sum_{s=0}^{p-1} \left(\sum_{l \in \mu^{-1}(s)} \zeta_l \right) = \sum_{s=0}^{p-1} \left(\sum_{l \in \mu^{-1}(s)} \tilde{\zeta}_l \right) \eta^s \tag{B.9}$$

We need the following algebraic fact (see e.g. [81], p. 60-61):

Proposition B.6. *The minimal polynomial of $\eta = e^{2\pi i/p^k}$ over the field $K = \mathbb{Q}(e^{2\pi i/M'})$ is given by $X^p - \eta^p$ if $k \geq 2$ and $X^{p-1} + X^{p-2} + \ldots + X + 1$ if $k = 1$.*

We now claim that M is square-free, or equivalently that for any prime power p^k dividing M,

$$k = 1. \tag{B.10}$$

Indeed, equation (B.9) shows that the minimal polynomial of ζ has degree at most $p-1$, which by Proposition B.6 is only satisfied in the case $k = 1$.

Further we claim that there exists $\sigma \in \mathbb{C} \setminus \{0\}$ so that

$$\sum_{l \in \mu^{-1}(s)} \tilde{\zeta}_l = \sigma \quad \forall\, 0 \leq s \leq p-1. \tag{B.11}$$

The existence of such a σ follows from Proposition B.6: As $k = 1$ by (B.10), the minimal polynomial of η over K is given by $X^{p-1} + X^{p-2} + \ldots + X + 1$. Since this is a polynomial of degree $p-1$ the polynomial on the right hand side of (B.9) must be a scalar multiple of the minimal polynomial. Hence all the coefficients $\sum_{l \in \mu^{-1}(s)} \tilde{\zeta}_l$ have the same value $\sigma \in \mathbb{C}$. As $\sum_{l \in \mu^{-1}(s)} \zeta_l = \sigma \eta^s$ the additional property $\sigma \neq 0$ follows from the assumption that there are no vanishing subsums. Hence we can assume w.l.o.g. that $\sigma = 1$.

Next we claim that

$$p \leq n. \tag{B.12}$$

In other words, possible prime factors of M are bounded by the number of summands in (B.7). To prove (B.12), note that it follows from (B.11) that for any $0 \leq s \leq p-1$ there exists $1 \leq l \leq n$ such that $\mu(l) = s$, i.e. the map $\mu : \{1, \ldots, n\} \to \{0, \ldots, p-1\}$ is onto. This establishes (B.12).

The map μ induces the *partition* $(\sharp\, \mu^{-1}(s))_{0 \leq s \leq p-1}$ of the positive integer n into p summands,

$$n = \sum_{s=0}^{p-1} \sharp\, \mu^{-1}(s) \tag{B.13}$$

Lemma B.7. *For any solution $\{\zeta_1, \dots, \zeta_8\}$ of (B.6) contained in S^1 without vanishing subsums there exists $\xi \in S^1$ such that either*

$$\{\zeta_1, \dots, \zeta_8\} = \{-\xi\alpha, -\xi\alpha^2\} \cup \{\xi\gamma^j \,|\, 1 \leq j \leq 6\} \tag{B.14}$$

or

$$\{\zeta_1, \dots, \zeta_8\} = \{-\xi\alpha^l, -\xi\alpha^l \cdot \beta^i, -\xi\alpha^l \cdot \beta^j \,|\, 1 \leq l \leq 2\} \cup \{\xi\beta^k, \xi\beta^m\}, \tag{B.15}$$

where the quadruple (i, j, k, m) is a permutation of $(1, 2, 3, 4)$ and

$$\alpha := e^{\frac{2\pi i}{3}}, \quad \beta := e^{\frac{2\pi i}{5}}, \quad \gamma := e^{\frac{2\pi i}{7}}.$$

Proof. By a straightforward computation one verifies that the sets of the form (B.14) or (B.15) satisfy (B.6). It remains to prove that these are the only solutions of (B.6) of this type.

We classify the solutions of (B.6) according to the possible values of p, which we now assume to be the largest prime dividing M. Since $n = 8$, by (B.12), the possible values of p are 2, 3, 5, and 7. If $p = 2$, then, by (B.10), $M = 2$ and therefore there exists $\xi \in S^1$ so that $\zeta_j = \pm\xi$ for any $1 \leq j \leq n$. In this case there exists a solution of (B.7) without vanishing subsums only if $n = 2$. (In this case, they are given by $\{\zeta_1, \zeta_2\} = \xi\{1, -1\}$ with $\xi \in S^1$.) If $p = 3$, then $M = 3$ or $M = 3 \cdot 2$, and there exists a solution of (B.7) without vanishing subsums only if $n = 3$. (In this case, they are given by $\{\zeta_1, \zeta_2, \zeta_3\} = \xi\{1, \alpha, \alpha^2\}$ with $\xi \in S^1$.) If $p = 5$, then $\eta = \beta$ in (B.8). Up to permutations, there are the following three partitions of 8 into 5 summands, $(4, 1, 1, 1, 1)$, $(3, 2, 1, 1, 1)$, and $(2, 2, 2, 1, 1)$. In a straightforward way one shows that the partitions $(4, 1, 1, 1, 1)$ and $(3, 2, 1, 1, 1)$ and their permutations give rise to solutions of the equation (B.6) with vanishing subsums. E.g. the solutions corresponding to $(4, 1, 1, 1, 1)$ are given by $\xi \cdot (-\beta, -\beta^2, -\beta^3, -\beta^4, \beta, \beta^2, \beta^3, \beta^4)$ with $\xi \in S^1$, whereas the solutions corresponding to $(3, 2, 1, 1, 1)$ are $\xi \cdot (-i, 1, i, -\alpha\beta, -\alpha^2\beta, \beta^2, \beta^3, \beta^4)$ with $\xi \in S^1$. On the other hand the partition $(2, 2, 2, 1, 1)$ leads to the solutions

$$(\zeta_1, \dots, \zeta_8) = \xi(-\alpha, -\alpha^2, -\alpha\beta, -\alpha^2\beta, -\alpha\beta^2, -\alpha^2\beta^2, \beta^3, \beta^4)$$

with $\xi \in S^1$. They are the solutions (B.15) with $(i, j, k, m) = (1, 2, 3, 4)$. Permutations of the partition $(2, 2, 2, 1, 1)$ again lead to solutions of the type (B.15), but with (i, j, k, m) given by a permutation of $(1, 2, 3, 4)$.

If $p = 7$, then $\eta = \gamma$ in (B.8). Then, up to permutations, $(2, 1, 1, 1, 1, 1, 1)$ is the only possible partition of 8 into 7 summands. The partition $(2, 1, 1, 1, 1, 1, 1)$ leads to the solutions

$$(\zeta_1, \dots, \zeta_8) = \xi(-\alpha, -\alpha^2, \gamma, \dots, \gamma^6)$$

with $\xi \in S^1$, where we used that $1 = -\alpha - \alpha^2$. They are of type (B.14). Any permutation of $(2, 1, 1, 1, 1, 1, 1)$ leads to the same kind of solutions. ∎

Lemma B.8. *For any solution $\{\zeta_1, \dots, \zeta_8\}$ of (B.6) contained in S^1 without vanishing subsums of length 2 but having a vanishing subsum of length 3, 4, or 5, there exist ξ, $\xi' \in S^1$ such that*

$$\{\zeta_1, \dots, \zeta_8\} = \{\xi\alpha^l | 0 \le l \le 2\} \cup \{\xi'\beta^m | 0 \le m \le 4\}, \tag{B.16}$$

where again $\alpha = e^{2\pi i/3}$ and $\beta = e^{2\pi i/5}$.

Proof. Again, one verifies by a direct computation that the solutions (B.16) of (B.6) have the desired properties. It remains to prove that they are the only ones. First note that under the hypotheses of the lemma, vanishing subsums of length 4 cannot occur, since the latter ones would imply the existence of vanishing subsums of length 2, which by assumption is excluded. Hence, in order to find solutions of (B.7) for $n = 8$ with the desired properties, we have to find all solutions of (B.7) without vanishing subsums for $n = 3$ and $n = 5$. Note that by (B.12), $p = n$ for $n = 3$ or $n = 5$. By the considerations in the proof of Lemma B.7, the former ones are given by $(\zeta_1, \zeta_2, \zeta_3) = \xi(1, \alpha, \alpha^2)$ and the latter ones by $(\zeta_1, \dots, \zeta_5) = \xi'(1, \beta, \beta^2, \beta^3, \beta^4)$ with $\xi, \xi' \in S^1$. This proves the lemma. □

We are now ready to prove Lemma B.5.

Proof of Lemma B.5. We first select from (B.14), (B.15) and (B.16) all the solutions $(\zeta_1, \dots, \zeta_8)$ of (B.6) which are of the form (B.3) (after multiplication by $2i$). This amounts to selecting the solutions $(\zeta_1, \dots, \zeta_8)$ of (B.6) having the property that $\{\zeta_1, \dots, \zeta_8\}$ is invariant under the map $\zeta \mapsto -\zeta^{-1}$. It requires to choose ξ and ξ' in (B.14), (B.15), and (B.16) appropriately. Let us explain this procedure in detail for the solutions of type (B.14).

First we rewrite the solution (B.14),

$$(\zeta_1, \dots, \zeta_8) = \xi \cdot (-\alpha, -\alpha^2, \gamma, \gamma^2, \gamma^3, \gamma^4, \gamma^5, \gamma^6) = e^{\frac{2\pi i x}{42}} \left(e^{\frac{2\pi i t_k}{42}}\right)_{1 \le k \le 8},$$

where $\xi = e^{2\pi i x/42}$ with $x \in \mathbb{R}/42\mathbb{Z}$ and

$$(t_1, \dots, t_8) = (6, 7, 12, 18, 24, 30, 35, 36). \tag{B.17}$$

The required invariance of the set of the ζ_k's under the map $\zeta \mapsto -\zeta^{-1}$ is equivalent to the invariance of the set of the $(t_k + x)$'s under the map $t \mapsto 21 - t$ (mod 42). Since the set (B.17) of the t_k's is invariant under the map $t \mapsto -t$ (mod 42), $\{t_k + x | 1 \le k \le 8\}$ is invariant under $t \mapsto 21 - t$ (mod 42), if we choose $x := \frac{21}{2}$ or $\xi = i$. Then the equation $\sum_{i=1}^{8} \zeta_i = 0$ reads

$$e^{\frac{11\pi i}{14}} + e^{\frac{5\pi i}{6}} + e^{\frac{15\pi i}{14}} + e^{\frac{19\pi i}{14}} + e^{\frac{23\pi i}{14}} + e^{\frac{27\pi i}{14}} + e^{\frac{\pi i}{6}} + e^{\frac{3\pi i}{14}} = 0,$$

or $\sin\frac{\pi}{6} + \sin\frac{3\pi}{14} + \sin\frac{15\pi}{14} + \sin\frac{19\pi}{14} = 0$. Choosing all arguments in $(0, \pi)$, the latter identity reads

$$\sin\frac{\pi}{6} + \sin\frac{3\pi}{14} - \sin\frac{\pi}{14} - \sin\frac{5\pi}{14} = 0. \tag{B.18}$$

For the solutions of type (B.15), one gets

$$\sin\frac{\pi}{6}+\sin\frac{13\pi}{30}-\sin\frac{7\pi}{30}-\sin\frac{3\pi}{10}=0 \tag{B.19}$$

and

$$\sin\frac{\pi}{6}+\sin\frac{\pi}{30}-\sin\frac{11\pi}{30}+\sin\frac{\pi}{10}=0. \tag{B.20}$$

Let us briefly explain how (B.19)-(B.20) can be obtained. Note that from the 24 permutations of $(1,2,3,4)$ in (B.15), there are only six which lead to different sets of the ζ_i's, since interchanging i and j or k and m leaves the set on the right hand side of (B.15) invariant. In the resulting six different cases, we again write $\{\zeta_1,\ldots,\zeta_8\}=\xi\cdot\{e^{2\pi i\cdot\frac{t_1}{30}},\ldots,e^{2\pi i\cdot\frac{t_8}{30}}\}$ with t_k in $\mathbb{R}/30\mathbb{Z}$. Then, up to translations, there are only two different types of solutions emerging from these six cases. With the appropriate choices of ξ, one gets the solutions (B.19) and (B.20).

Finally, for the solutions of type (B.16), one gets

$$\sin\frac{\pi}{2}-\sin\frac{\pi}{6}+\sin\frac{\pi}{10}-\sin\frac{3\pi}{10}=0. \tag{B.21}$$

The procedure to obtain (B.21) is basically the same as in the preceding cases. We write (B.16) as $\{\zeta_1,\ldots,\zeta_8\}=\xi\cdot\{\alpha^l,\ \lambda\cdot\beta^m|0\le l\le 2,\ 0\le m\le 4\}$ and first choose $\lambda\in S^1$ so that the set $\{\alpha^l,\ \lambda\cdot\beta^m|0\le l\le 2,\ 0\le m\le 4\}$ is symmetric with respect to some axis through the origin, and then choose ξ so that this axis is the imaginary axis.

To finish the proof of Lemma B.5 it remains to show that all the solutions (k_1,k_2,k_3,k_4) of $\sum_{i=1}^4\sin\frac{k_i\pi}{N}=0$ obtained in (B.18)-(B.21) and the additional ones obtained by replacing $0<x<\pi$ in $\sin x$ by $\pi-x$ satisfy $\sum_{i=1}^4 k_i\not\equiv 0 \mod N$ and hence are not in K_4.

For the solutions obtained in (B.18)-(B.21), N is even. Hence if N is odd, then there is no quadruple $(k_1,k_2,k_3,k_4)\in K_4$ such that (B.3) and (B.4) are satisfied. This finishes the proof of Lemma B.5 in this case.

For the rest of the proof, we assume that N is even. If $N=42r$ for some $r\in\mathbb{N}$, (B.18) becomes

$$\sin\frac{7r\pi}{42r}+\sin\frac{9r\pi}{42r}+\sin\frac{(-3r)\pi}{42r}+\sin\frac{(-15r)\pi}{42r}=0,$$

and we have $7r+9r-3r-15r=-2r\not\equiv 0 \mod 42r$. Hence the corresponding quadruple (k_1,k_2,k_3,k_4) is not in K_4. For the quadruples obtained by replacing $0<x<\pi$ in $\sin x$ by $\pi-x$ in some of the summands in (B.18), the condition $\sum_{i=1}^4 k_i\not\equiv 0 \mod 42r$ amounts to

$$\pm 7\pm 9\pm 3\pm 15\not\equiv 0 \mod 42 \tag{B.22}$$

for any combination of plus and minus signs. The relations (B.22) are easily verified. Similarly, one verifies that the quadruples (k_1,k_2,k_3,k_4) satisfying (B.19), (B.20), or (B.21) are not in K_4 by showing that

$$\pm 5\pm 13\pm 7\pm 9\not\equiv 0,\quad \pm 5\pm 1\pm 11\pm 3\not\equiv 0,\quad \pm 15\pm 5\pm 3\pm 9\not\equiv 0 \mod 30, \tag{B.23}$$

again for any combination of plus and minus signs. Hence we have shown that none of the solutions (k_1, k_2, k_3, k_4) of (B.3) is an element of K_4. This completes the proof of Lemma B.5. $\square$

Proof of Lemma B.1. The claimed statement follows from the Lemmas B.2, B.3, B.4, and B.5. $\square$

Appendix C

Proof of Lemma 3.3.1

In order to prove Lemma 3.3.1, we need to express the map S, defined by (3.63), with respect to the coordinates $(x_k, y_k)_{1\le k\le N-1}$ of Theorem 1.3.1. The transformation, defined on a neighborhood of $0 \in \mathcal{Z}$ in section 3.1,

$$(x, y) = (x_k, y_k)_{1\le k\le N-1} \mapsto (q, p) = (q_n, p_n)_{1\le n\le N} \in \mathcal{M},$$

is given by the composition $\Psi = \Psi_0 \circ \Psi_1 \circ \Psi_2$ of the canonical transformations Ψ_0, Ψ_1, and Ψ_2. Let $(q, p) = \Psi_0(\zeta^{(2)})$, $\zeta^{(2)} = \Psi_1(\zeta^{(3)})$, and $\zeta^{(3)} = \Psi_2(\zeta^{(4)})$, where $\zeta^{(4)} \equiv \zeta$ are the complex coordinates related to $(x_k, y_k)_{1\le k\le N-1}$ by (3.65). Note that Ψ_0 thus is the composition of the transformations $\zeta \mapsto (v, u)$ and $(v, u) \mapsto (q, p)$, where $(v, u) = (v_i, u_i)_{1\le i\le N-1}$ are the relative coordinates introduced by (3.1). Recall from section 3.1 that $(q, p) = \Psi_0(\zeta^{(2)}) \in \mathcal{M}$ is the linear transformation given by

$$-p_n + P = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \sqrt{|s_k|}e^{\pi i(2n-2)k/N}\zeta_k^{(2)} \quad (1 \le n \le N-1), \tag{C.1}$$

$$q_{n+1} - q_n = \frac{1}{\sqrt{N}} \sum_{1\le|k|\le N-1} \sqrt{|s_k|}e^{\pi i(2n-1)k/N}\zeta_k^{(2)} \quad (1 \le n \le N-1), \tag{C.2}$$

where $P = \frac{1}{N}\sum_{n=1}^{N} p_n$ is the total momentum of the chain. For simplicity we assume in the sequel that $P = 0$; the general case is completely analogous (in the following application to $\mathrm{Fix}(S)$, we indeed have $P = 0$). Note that (C.1) and (C.2) continue to hold for $n = N$. For example, to see this for (C.2), we write

$$-(q_1 - q_N) = \sum_{k=1}^{N-1}(q_{k+1} - q_k)$$

and substitute the expressions (C.2) for $q_{k+1} - q_k$ into the latter sum, from which the claim follows.

Solving (C.1)-(C.2) for $\zeta_k^{(2)}$ $(1 \le |k| \le N-1)$ one gets for $\zeta_k^{(2)} \equiv (\Psi_0^{-1}(q,p))_k$

$$\begin{aligned}
\zeta_k^{(2)} &= \frac{1}{2\sqrt{N|s_k|}} \sum_{n=1}^{N} \Big(-e^{-\pi i(2n-2)k/N} p_n + e^{-\pi i(2n-1)k/N} (q_{n+1} - q_n) \Big) \\
&= \frac{1}{2\sqrt{N|s_k|}} \sum_{n=1}^{N} \Big(-e^{-\pi i(2n-2)k/N} p_n + (e^{-\pi i(2n-3)k/N} - e^{-\pi i(2n-1)k/N}) q_n \Big) \\
&= \frac{1}{2\sqrt{N|s_k|}} \sum_{n=1}^{N} e^{-\pi i(2n-2)k/N} \Big(-p_n + (e^{\pi ik/N} - e^{-\pi ik/N}) q_n \Big) \\
&= \frac{e^{2\pi ik/N}}{2\sqrt{N|s_k|}} \sum_{n=1}^{N} e^{-2\pi ink/N} (-p_n + 2is_k q_n). \qquad \text{(C.3)}
\end{aligned}$$

The transformations Ψ_1 and Ψ_2 are defined locally around the origin of $\mathcal{Z}$ and are given by $\Psi_1 = X_{F_3}^t|_{t=1}$ and $\Psi_2 = X_{F_4}^t|_{t=1}$, where the Hamiltonians F_3 and F_4 are homogeneous polynomials of order three respectively four in $(\zeta_k)_{1 \le |k| \le N-1}$. Inverting Ψ_1 and Ψ_2, we obtain

$$\zeta^{(3)} = (X_{F_3}^t|_{t=-1})(\zeta^{(2)}) \quad \text{and} \quad \zeta^{(4)} = (X_{F_4}^t|_{t=-1})(\zeta^{(3)}). \qquad \text{(C.4)}$$

Proof of Lemma 3.3.1. Recall first that $\mathrm{Fix}(S)$, defined in terms of the coordinates (q,p) by (3.64), is a symplectic submanifold of $\mathcal{M} \subseteq T^*\mathbb{R}^N$, and that all three transformations Ψ_0, Ψ_1, and Ψ_2 are canonical. We first show that $S_{\mathcal{Z}} \circ \Psi_0^{-1} = \Psi_0^{-1} \circ S$, and then that $S_{\mathcal{Z}}$ commutes with Ψ_1 and Ψ_2 (cf. (3.66) for the definition of $S_{\mathcal{Z}}$). This then shows the claimed identity $S \circ \Psi = \Psi \circ S_{\mathcal{Z}}$.

To prove that $S_{\mathcal{Z}} \circ \Psi_0^{-1} = \Psi_0^{-1} \circ S$, we compute for any $(q,p) \in \mathrm{Fix}(S)$ and any $1 \le k \le N-1$

$$\begin{aligned}
(S_{\mathcal{Z}}(\Psi_0^{-1}(q,p)))_k &= -e^{4\pi ik/N} (\Psi_0^{-1}(q,p))_{N-k} \\
&= -e^{4\pi ik/N} \frac{e^{-2\pi ik/N}}{2\sqrt{N|s_{N-k}|}} \sum_{l=1}^{N} e^{-2\pi il(N-k)/N} (-p_l + 2is_{N-k} q_l) \\
&= \frac{e^{2\pi ik/N}}{2\sqrt{N|s_k|}} \sum_{l=1}^{N} e^{2\pi ilk/N} (p_l - 2is_k q_l),
\end{aligned}$$

using the formulas (3.66) and (C.3) for $S_{\mathcal{Z}}$ and Ψ_0^{-1} in the first and second equalities, respectively, and the identity $s_{N-k} = s_k$ in the third equality. Note that the (vanishing) summand $l = N$ can be omitted in the latter sum, in which we now substitute l by $N - l$, obtaining

$$\begin{aligned}
(S_{\mathcal{Z}}(\Psi_0^{-1}(q,p)))_k &= \frac{e^{2\pi ik/N}}{2\sqrt{N|s_k|}} \sum_{l=1}^{N-1} e^{2\pi i(N-l)k/N} (p_{N-l} - 2is_k q_{N-l}) \\
&= \frac{e^{2\pi ik/N}}{2\sqrt{N|s_k|}} \sum_{l=1}^{N} e^{-2\pi ilk/N} (-\tilde{S}(p)_l + 2is_k \tilde{S}(q)_l) \\
&= (\Psi_0^{-1}(S(q,p)))_k.
\end{aligned}$$

In the second equality, we have again included the (vanishing) summand $l = N$. Further we write $S(q,p) = (\tilde{S}(q), \tilde{S}(p))$. In particular we have shown that $\mathrm{Fix}(S_{\mathcal{Z}})$ is the image of $\mathrm{Fix}(S)$ under Ψ_0^{-1}.

Next we claim that $F_3 : \mathcal{Z} \to \mathbb{C}$ is invariant under $S_{\mathcal{Z}}$, $F_3 \circ S_{\mathcal{Z}} = F_3$. Recall from (3.17) and (3.22) that F_3 is given by the polynomial

$$F_3 = \sum_{(k,k',k'')\in K_3} F^{(3)}_{kk'k''} \zeta_k \zeta_{k'} \zeta_{k''},$$

with K_3 denoting the set of all $(k,k',k'') \in \mathbb{Z}^3$ satisfying $1 \le |k|, |k'|, |k''| \le N-1$ and $k + k' + k'' \equiv 0 \bmod N$, and

$$F^{(3)}_{kk'k''} = \frac{(-1)^{(k+k'+k'')/N}}{12i\sqrt{N}} \frac{\sqrt{|s_k s_{k'} s_{k''}|}}{s_k + s_{k'} + s_{k''}}.$$

As $s_{N-k} = s_k$ one has $F^{(3)}_{N-k,N-k',N-k''} = -F^{(3)}_{kk'k''}$, and $(k,k',k'') \in K_3$ if and only if $(N-k, N-k', N-k'') \in K_3$. Here we view $N-k$, $N-k'$, and $N-k''$ mod $2N$ and replace them if necessary by representatives in $\{\pm 1, \ldots, \pm(N-1)\}$. Hence

$$\begin{aligned}
F_3(S_{\mathcal{Z}}(\zeta)) &= \sum_{(k,k',k'')\in K_3} F^{(3)}_{kk'k''} (S_{\mathcal{Z}}(\zeta))_k (S_{\mathcal{Z}}(\zeta))_{k'} (S_{\mathcal{Z}}(\zeta))_{k''} \\
&= \sum_{(k,k',k'')\in K_3} F^{(3)}_{kk'k''} (-1)^3 e^{-4\pi i (k+k'+k'')/N} \zeta_{N-k} \zeta_{N-k'} \zeta_{N-k''} \\
&= -\sum_{(k,k',k'')\in K_3} F^{(3)}_{kk'k''} \zeta_{N-k} \zeta_{N-k'} \zeta_{N-k''} \\
&= -\sum_{(k,k',k'')\in K_3} F^{(3)}_{N-k,N-k',N-k''} \zeta_k \zeta_{k'} \zeta_{k''} \\
&= \sum_{(k,k',k'')\in K_3} F^{(3)}_{kk'k''} \zeta_k \zeta_{k'} \zeta_{k''} = F_3(\zeta).
\end{aligned}$$

From $F_3 \circ S_{\mathcal{Z}} = F_3$ it follows (by considering a Taylor expansion of Ψ_1) that near $0 \in \mathcal{Z}$ where Ψ_1 is defined, $\Psi_1 = X^t_{F_3}|_{t=1}$ commutes with $S_{\mathcal{Z}}$. In particular, $\mathrm{Fix}(S_{\mathcal{Z}})$ is invariant under the flow $X^t_{F_3}$.

It remains to show that near $0 \in \mathcal{Z}$ where Ψ_2 is defined, $\Psi_2 = X^t_{F_4}|_{t=1}$ commutes with $S_{\mathcal{Z}}$. As above, this follows from $F_4 \circ S_{\mathcal{Z}} = F_4$. Recall from (3.35) and (3.38) that F_4 is given by the polynomial

$$F_4 = \sum_{(k,k',k'',k''')\in K_4 \setminus K_4^N} F^{(4)}_{kk'k''k'''} \zeta_k \zeta_{k'} \zeta_{k''} \zeta_{k'''}.$$

Here $K_4 \setminus K_4^N$ denotes the set of all quadruples $(k,k',k'',k''') \in \mathbb{Z}^4$ satisfying $1 \le |k|, |k'|, |k''|, |k'''| \le N-1$ and $k + k' + k'' + k''' \equiv 0 \bmod N$ such that there do not exist $1 \le l \le m \le N-1$ with $\{k,k',k'',k'''\} = \{l, m, -l, -m\}$, and

$$F^{(4)}_{kk'k''k'''} = \frac{(-1)^{(k+k'+k''+k''')/N}}{24i\sqrt{N}} \frac{(\beta + \frac{3\alpha^2}{2} c^S_{kk'k''k'''}) \sqrt{|s_k s_{k'} s_{k''} s_{k'''}|}}{s_k + s_{k'} + s_{k''} + s_{k'''}}, \tag{C.5}$$

where $c^S_{kk'k''k'''} := \frac{1}{4!}\sum_{\sigma\in S_4} c_{\sigma(k,k',k'',k''')}$ and

$$c_{lml'm'} = \begin{cases} \frac{1}{-1+\frac{s_{l'}+s_{m'}}{|s_{l'+m'}|}} - \frac{1}{1+\frac{s_l+s_m}{|s_{l+m}|}} & \text{if } l+m \not\equiv 0 \bmod N \\ 0 & \text{otherwise.} \end{cases} \tag{C.6}$$

First note that $(k,k',k'',k''') \in K_4\setminus K_4^N$ if and only if $(N-k, N-k', N-k'', N-k''') \in K_4 \setminus K_4^N$ (where here we again view $N-k$, $\ldots$, $N-k'''$ mod $2N$ and replace them if necessary by representatives in $\{\pm 1, \ldots, \pm(N-1)\}$). Next, it follows from the definition (C.6) of $c_{kk'k''k'''}$ and the identities $s_{N-k} = s_k$ that $c_{N-k,N-k',N-k'',N-k'''} = c_{kk'k''k'''}$. Hence $c^S_{N-k,N-k',N-k'',N-k'''} = c^S_{kk'k''k'''}$ and, by the definition (C.5) of $F^{(4)}_{kk'k''k'''}$, $F^{(4)}_{N-k,N-k',N-k'',N-k'''} = F^{(4)}_{kk'k''k'''}$. Thus

$$\begin{aligned} F_4(S_{\mathcal{Z}}(\zeta)) &= \sum_{(k,k',k'',k''')\in K_4\setminus K_4^N} F^{(4)}_{kk'k''k'''}(S_{\mathcal{Z}}(\zeta))_k (S_{\mathcal{Z}}(\zeta))_{k'} (S_{\mathcal{Z}}(\zeta))_{k''} (S_{\mathcal{Z}}(\zeta))_{k'''} \\ &= \sum_{(k,k',k'',k''')\in K_4\setminus K_4^N} F^{(4)}_{kk'k''k'''}(-1)^4 e^{-4\pi i(k+k'+k''+k''')/N} \zeta_{N-k}\zeta_{N-k'}\zeta_{N-k''}\zeta_{N-k'''} \\ &= \sum_{(k,k',k'',k''')\in K_4\setminus K_4^N} F^{(4)}_{kk'k''k'''}\zeta_{N-k}\zeta_{N-k'}\zeta_{N-k''}\zeta_{N-k'''} \\ &= \sum_{(k,k',k'',k''')\in K_4\setminus K_4^N} F^{(4)}_{N-k,N-k',N-k'',N-k'''}\zeta_k\zeta_{k'}\zeta_{k''}\zeta_{k'''} \\ &= F_4(\zeta). \end{aligned}$$

This proves $F_4 \circ S_{\mathcal{Z}} = F_4$. Therefore (again by considering a Taylor expansion of Ψ_2), near $0 \in \mathcal{Z}$ where Ψ_2 is defined, $\Psi_2 = X^t_{F_4}|_{t=1}$ commutes with $S_{\mathcal{Z}}$. $\square$

Appendix D

Proof of Lemma 4.1.5

Here we prove Lemma 4.1.5 on the values of the N basic symmetric polynomials in $N-1$ variables. These polynomials are given by

$$\Pi_0 := 1, \tag{D.1}$$

$$\Pi_n(t_1, \ldots, t_{N-1}) := \sum_{1 \le i_1 < \ldots < i_n \le N-1} t_{i_1} \cdot \ldots \cdot t_{i_n} \ (1 \le n \le N-1). \tag{D.2}$$

The lemma on these polynomials we prove in this appendix is the following one.

Lemma D.1. *Let $N \ge 2$ be an arbitrary integer (not necessarily odd). Evaluated at $t_k = \sin^2 \frac{k\pi}{N}$ $(1 \le k \le N-1)$, the N basic symmetric polynomials $(\Pi_n)_{0 \le n \le N-1}$ given by (D.1) and (D.2) take the values*

$$\Pi_n\left(\sin^2 \frac{\pi}{N}, \ldots, \sin^2 \frac{(N-1)\pi}{N}\right) = 4^{-n} \frac{N}{N-n} \binom{2N-n-1}{n}. \tag{D.3}$$

Proof. For fixed $N \ge 2$, we proceed by induction on n. For $n = 0, 1$, (D.3) is easily verified. For the induction step, we use the formulas of Newton-Girard (see e.g. [76], p. 278-279) expressing for arbitrary $t_1, \ldots, t_{N-1}$ the polynomials $\Pi_n(t_1, \ldots, t_{N-1})$ inductively through the Newton sums

$$S_n(t_1, \ldots, t_{N-1}) := \sum_{k=1}^{N-1} t_k^n \tag{D.4}$$

by (omitting the arguments)

$$m\,\Pi_m + \sum_{k=1}^{m} (-1)^k S_k \Pi_{m-k} = 0, \tag{D.5}$$

for any $1 \le m \le N-1$. To apply (D.5), we first cite from ([67], p. 640) the value of the Newton sums (D.4) for $t_k = \sin^2 \frac{k\pi}{N}$, namely

$$S_n\left(\sin^2 \frac{\pi}{N}, \ldots, \sin^2 \frac{(N-1)\pi}{N}\right) = \sum_{k=1}^{N-1} \sin^{2n} \frac{k\pi}{N} = \frac{N}{4^n} \binom{2n}{n}. \tag{D.6}$$

Assume now that (D.3) is verified for all $0 \leq m \leq n$ and that $n+1 \leq N-1$. Use the identity

$$\frac{1}{r}\binom{r}{l} = \frac{1}{r-l}\binom{r-1}{l} \quad (r, l \in \mathbb{N},\ l < r) \tag{D.7}$$

for $r := 2N-n$, $l := n$ to observe that the right hand side of (D.3) is equal to $4^{-n}\frac{2N}{2N-n}\binom{2N-n}{n}$. Substitute (D.3) and (D.6) into the Newton-Girard formula (D.5) with $m := n+1$, i.e.

$$\Pi_{n+1} = -\frac{1}{n+1}\sum_{k=1}^{n+1}(-1)^k S_k \Pi_{n+1-k},$$

to get for $\Pi_{n+1} = \Pi_{n+1}\big(\sin^2\frac{\pi}{N}, \ldots, \sin^2\frac{(N-1)\pi}{N}\big)$

$$\begin{aligned}
\Pi_{n+1} &= -\frac{1}{n+1}\sum_{k=1}^{n+1}(-1)^k \frac{N\cdot 4^{-(k+(n+1-k))}\cdot 2N}{2N-(n+1-k)}\binom{2k}{k}\binom{2N-(n+1-k)}{n+1-k} \\
&= -\frac{2N^2\cdot 4^{-(n+1)}}{n+1}\sum_{k=1}^{n+1}\frac{(-1)^k}{2N-(n+1-k)}\binom{2k}{k}\binom{2N-(n+1-k)}{n+1-k} \\
&\overset{*}{=} 4^{-(n+1)}\frac{N}{N-(n+1)}\binom{2N-(n+1)-1}{n+1}.
\end{aligned}$$

The equality (*) remains to be proven for the induction step to be complete. Writing m instead of $n+1$, the proof of (*) amounts to proving that for any $N \geq 2$ and $1 \leq m \leq N-1$,

$$\sum_{k=1}^{m}\frac{(-1)^k}{2N-(m-k)}\binom{2k}{k}\binom{2N-(m-k)}{m-k} = -\frac{m}{2N(N-m)}\binom{2N-m-1}{m},$$

or, including $k=0$ in the summation, again using (D.7) for $r := 2N-(m-k)$ and $l := m-k$ and multiplying by 2,

$$\sum_{k=0}^{m}\frac{(-1)^k}{N-(m-k)}\binom{2k}{k}\binom{2N-(m-k)-1}{m-k} = \frac{1}{N}\binom{2N-m-1}{m}. \tag{D.8}$$

Set $N' := N-m$ (after which the $'$ is again omitted); then (D.8) is equivalent to claiming that for any $N, m \in \mathbb{N}$,

$$\sum_{k=0}^{m}\frac{(-1)^k}{N+k}\binom{2k}{k}\binom{2N+m+k-1}{m-k} = \frac{1}{N+m}\binom{2N+m-1}{m}. \tag{D.9}$$

To prove (D.9), we proceed inductively in m for any fixed N. First note that the identity holds for $m=0$ and $m=1$, since in these cases both sides of (D.9) are equal to $\frac{1}{N}$ and $\frac{2N}{N+1}$, respectively. For the induction step, we claim that both sides of (D.9) satisfy the recurrence relation

$$(N+m)(2N+m)f_m - (m+1)(N+m+1)f_{m+1} = 0 \tag{D.10}$$

for any $m \in \mathbb{N}$. That the right hand side of (D.9) satisfies (D.10) can be checked by a direct computation. To see that the same holds for the left hand side of (D.9), first note that the summand

$$F_{k,m} := \frac{(-1)^k}{N+k}\binom{2k}{k}\binom{2N+m+k-1}{m-k}$$

of the left hand side of (D.9) satisfies the recurrence relation

$$(N+m)(2N+m)F_{k,m} - (m+1)(N+m+1)F_{k,m+1} = \Delta_k(F_{k,m}R_{k,m}), \quad \text{(D.11)}$$

where for $k \le m$

$$R_{k,m} := \frac{k(N+k)(2N+2k-1)}{m-k+1},$$

and where Δ_k is the standard discrete difference operator of order one in k, i.e. $\Delta_k(F_{k,m}R_{k,m}) = F_{k+1,m}R_{k+1,m} - F_{k,m}R_{k,m}$. The proof of (D.11) is a straightforward (lengthy) computation.

The recurrence relations (D.10) and (D.11) were found using the Mathematica program `zb-alg.m` written by P. Paule, M. Schorn & A. Riese, which is an implementation of D. Zeilberger's "creative telescoping" algorithm, described e.g. in Chapter 6 of [60]. We however emphasize that the *proof* of both recurrence relations is purely analytical, the program just mentioned was "only" used to *find* the recurrence relations.

Having proved that $F_{k,m}$ satisfies (D.11), we return to the proof of the fact that the left hand side of (D.9) satisfies the recurrence relation (D.10). Denote the left hand side of (D.9) by f_m, i.e. for any $m \in \mathbb{N}$

$$f_m := \sum_{k=0}^{m} F_{k,m} = \sum_{k=0}^{m} \frac{(-1)^k}{N+k}\binom{2k}{k}\binom{2N+m+k-1}{m-k}.$$

We sum the left hand side of (D.11) for k from 0 to $m-1$ and obtain

$$\begin{aligned}(N+m)(2N+m)(f_m - F_{m,m}) - (m+1)(N+m+1)(f_{m+1} - (F_{m,m+1} + F_{m+1,m+1}))&\\ = F_{m,m}R_{m,m} - F_{0,m}R_{0,m}&\\ = (-1)^m m(2N+2m-1)\binom{2m}{m},& \quad \text{(D.12)}\end{aligned}$$

using (D.11) for the first and a direct computation for the second equality. Another (lengthy) computation shows that

$$\begin{aligned}(N+m)(2N+m)F_{m,m} - (m+1)(N+m+1)(F_{m,m+1} + F_{m+1,m+1})&\\ = -(-1)^m m(2N+2m-1)\binom{2m}{m}.&\end{aligned}$$

Adding this to (D.12) yields (D.10) and hence completes the proof of (D.9). As mentioned above, this completes the proof of Lemma D.1. □

Appendix E

Spectrum of the matrix P^D

Here we compute analytically for any integer $N' \geq 3$ the eigenvalues of the $N' \times N'$ matrix P^D, given by

$$\underbrace{\begin{pmatrix} 3 & 4 & \dots & & & & \dots & 4 & 2 \\ 4 & 3 & 4 & \dots & & \dots & 4 & 2 & 4 \\ \vdots & & \ddots & & & & \ddots & & \vdots \\ & & & 3 & 4 & 4 & 2 & & \\ & & & 4 & 3 & 2 & 4 & & \\ & & & 4 & 2 & 3 & 4 & & \\ & & & 2 & 4 & 4 & 3 & & \\ \vdots & & \ddots & & & & \ddots & & \vdots \\ 4 & 2 & 4 & \dots & & \dots & 4 & 3 & 4 \\ 2 & 4 & \dots & & & & \dots & 4 & 3 \end{pmatrix}}_{(N' \text{ even})}, \quad \underbrace{\begin{pmatrix} 3 & 4 & \dots & & & & \dots & 4 & 2 \\ 4 & 3 & 4 & \dots & & \dots & 4 & 2 & 4 \\ \vdots & & \ddots & & & & \ddots & & \vdots \\ & & & 3 & 4 & 2 & & & \\ & & & 4 & 2 & 4 & & & \\ & & & 2 & 4 & 3 & & & \\ \vdots & & \ddots & & & & \ddots & & \vdots \\ 4 & 2 & 4 & \dots & & \dots & 4 & 3 & 4 \\ 2 & 4 & \dots & & & & \dots & 4 & 3 \end{pmatrix}}_{(N' \text{ odd})}.$$

Lemma E.1. *If N' is even, the eigenvalues of P^D are $4N'-3$ (with multiplicity one), 1 (with multiplicity $\frac{N'}{2}$), and -3 (with multiplicity $\frac{N'}{2}-1$). If N' is odd, the eigenvalues of P^D are 1 (with multiplicity $\frac{N'-1}{2}$), -3 (with multiplicity $\frac{N'-3}{2}$), and $\frac{1}{2}(4N'-5)\left(1 \pm \sqrt{1+\frac{8(4N'-1)}{(4N'-5)^2}}\right)$ (each with multiplicity one).*

Proof. Throughout this proof, $\operatorname{antidiag}(a_1, \dots, a_{N'})$ denotes the "antidiagonal" $N' \times N'$-matrix M with $M_{kl} = a_l$ if $k + l = N' + 1$ and $M_{kl} = 0$ otherwise.

First consider the case where N' is even. We write P^D in the form

$$P^D = \operatorname{diag}(-1, \dots, -1) + \operatorname{antidiag}(-2, \dots, -2) + 4 \cdot 1_{N' \times N'}$$

and, with $\mu := -1 - \lambda$,

$$P^D - \lambda \mathrm{Id} = \underbrace{\operatorname{diag}(\mu, \dots, \mu) + \operatorname{antidiag}(-2, \dots, -2)}_{=:L^{(N')}} + 4 \cdot 1_{N' \times N'}.$$

Here $1_{N'\times N'}$ denotes the $N'\times N'$-matrix whose entries are all equal to 1.

We compute $\det(P^D - \lambda \mathrm{Id}) = \det(L^{(N')} + 4\cdot 1_{N'\times N'})$ by column expansion. Note that in the column expansion of the determinant only those terms contribute which are determinants of matrices containing at most one column consisting of entries all equal to four. We obtain

$$\det(P^D - \lambda \mathrm{Id}) = \det(L^{(N')}) + \sum_{j=1}^{N'} \det(L_j^{(N')}), \tag{E.1}$$

where $L_j^{(N')}$ is defined as the matrix $L^{(N')}$ with the j-th column replaced by the column $4\cdot(1,\dots,1)$. By expansion with respect to the first column and then the last column, the determinant of $L^{(N')}$ can be computed recursively,

$$\det(L^{(N')}) = (\mu^2 - 2^2)\det(L^{(N'-2)})$$

Since $\det(L^{(2)}) = \mu^2 - 4$, it follows by induction that

$$\det(L^{(N')}) = (\mu^2-4)^{\frac{N'}{2}}. \tag{E.2}$$

To compute $\det(L_1^{(N')})$, we expand the determinant in the same way and obtain the identity $\det(L_1^{(N')}) = 4(\mu+2)\det(L^{(N'-2)})$, from which it follows that

$$\det(L_1^{(N')}) = 4(\mu+2)(\mu^2-4)^{\frac{N'}{2}-1}. \tag{E.3}$$

Similarly one gets $\det(L_2^{(N')}) = (\mu^2-4)\det(L_1^{(N'-2)})$, and thus

$$\det(L_2^{(N')}) = \det(L_1^{(N')}) = 4(\mu+2)(\mu^2-4)^{\frac{N'}{2}-1}. \tag{E.4}$$

For any $1 < j < \frac{N'}{2}$, this procedure leads to $\det(L_j^{(N')}) = (\mu^2-4)\det(L_{j-1}^{(N'-2)})$ and hence

$$\det(L_j^{(N')}) = \det(L_1^{(N')}) = 4(\mu+2)(\mu^2-4)^{\frac{N'}{2}-1}. \tag{E.5}$$

For $j > \frac{N'}{2}$, note that $\det L_j^{(N')} = \det L_{N-j}^{(N')}$, since $L_j^{(N')}$ and $L_{N-j}^{(N')}$ can be transformed into each other by exchanging the j'th and the $(N'-j+1)$'th columns and then the j'th and the $(N'-j+1)$'th rows. By (E.2)-(E.5), we obtain

$$\begin{aligned}\det(P^D - \lambda \mathrm{Id}) &= (\mu^2-4)^{\frac{N'}{2}-1}\cdot\left((\mu^2-4) + N'\cdot 4(\mu+2)\right)\\ &= (\mu^2-4)^{\frac{N'}{2}-1}(\mu+2)(\mu-2+4N').\end{aligned}$$

Hence, if N' is *even*, the zeroes of $\det(P^D - \lambda \mathrm{Id})$ are $\mu = 2$ (with multiplicity $\frac{N'}{2}-1$), $\mu = -2$ (with multiplicity $\frac{N'}{2}$), and $\mu = -4N'+2$ (with multiplicity 1). Transforming back to $\lambda = -1-\mu$, we obtain the claimed eigenvalues.

It remains to consider the case where N' is *odd*. Again, we write

$$P^D = \mathrm{diag}(-1,\dots,-1,\overbrace{0}^{(\frac{N'+1}{2})},-1,\dots,-1) + \mathrm{antidiag}(-2,\dots,-2) + 4\cdot 1_{N'\times N'}.$$

With $\mu = -1 - \lambda$ we get

$$P^D - \lambda \mathrm{Id} = L^{(N')} + 4 \cdot 1_{N' \times N'}$$

with

$$L^{(N')} = \begin{pmatrix} \mu & 0 & \dots & 0 & -2 \\ 0 & \ddots & & \ddots & 0 \\ \vdots & & \mu - 1 & & \vdots \\ 0 & \ddots & & \ddots & 0 \\ -2 & 0 & \dots & 0 & \mu \end{pmatrix}.$$

As above, we obtain the expansion (E.1) for the determinant of $P^D - \lambda\mathrm{Id}$. We expand $\det(L^{(N')})$ with respect to the first column and then the last column, yielding the recursion formula

$$\det(L^{(N')}) = (\mu^2 - 4)\det(L^{(N'-2)}),$$

which together with $\det(L^{(1)}) = \mu - 1$ leads to

$$\det(L^{(N')}) = (\mu^2 - 4)^{\frac{N'-1}{2}}(\mu - 1). \tag{E.6}$$

For $\det(L_1^{(N')})$, we obtain the identity $\det(L_1^{(N')}) = 4(\mu + 2)\det(L^{(N'-2)})$ and hence

$$\det(L_1^{(N')}) = 4(\mu + 2)(\mu^2 - 4)^{\frac{N'-3}{2}}(\mu - 1). \tag{E.7}$$

More generally, for any $1 < j < \frac{N'}{2}$, we have

$$\det L_{N-j}^{(N')} = \det(L_j^{(N')}) = \det(L_1^{(N')}). \tag{E.8}$$

It remains to compute $\det L_{\frac{N'+1}{2}}^{(N')}$. Expanding $\det L_{\frac{N'+1}{2}}^{(N')}$ by the first column and then the last column, we obtain the recursion relation

$$\det L_{\frac{N'+1}{2}}^{(N')} = (\mu^2 - 4)\det L_{\frac{(N'-2)+1}{2}}^{(N'-2)}.$$

Together with $\det L_2^{(3)} = \det \begin{pmatrix} \mu & 4 & -2 \\ 0 & 4 & 0 \\ -2 & 4 & \mu \end{pmatrix} = 4(\mu^2 - 4)$, this implies

$$\det L_{\frac{N'+1}{2}}^{(N')} = 4(\mu^2 - 4)^{\frac{N'-1}{2}}. \tag{E.9}$$

Hence, combining (E.6)-(E.9) we obtain

$$\begin{aligned} \det(P^D - \lambda\mathrm{Id}) &= (\mu^2 - 4)^{\frac{N'-3}{2}} \cdot \big((\mu^2 - 4)(\mu - 1) + (N' - 1) \cdot 4(\mu + 2)(\mu - 1) + 4(\mu^2 - 4)\big) \\ &= (\mu^2 - 4)^{\frac{N'-3}{2}}(\mu + 2)(\mu^2 + (4N' - 3)\mu - (4N' + 2)). \end{aligned}$$

Hence, if N' is odd, the zeroes of $\det(P^D - \lambda \mathrm{Id})$ are $\mu = 2$ (with multiplicity $\frac{N'-3}{2}$), $\mu = -2$ (with multiplicity $\frac{N'-1}{2}$), and

$$\mu = -\frac{1}{2}(4N' - 3) \pm \frac{1}{2}\sqrt{16N'^2 - 8N' + 17}$$

(each with multiplicity 1). Transforming back to $\lambda = -1 - \mu$, we obtain the claimed formulas for the eigenvalues in the case where N' is odd. This completes the proof of Lemma E.1. $\square$

Appendix F

Critical points of $\mathcal{M}_\gamma$ of rank 0

In this appendix we study the nature of the critical points $\varepsilon(0,0,r_1,0,0,\pm r_2)$ of rank 0 of the map $\mathcal{M}_\gamma$, when viewed as critical points of K_γ. Recall that $\mathcal{M}_\gamma$ is the reduced moment map introduced in section 5.3. Throughout this appendix we use the notation of section 5.3 without any further comment. Choose $(M_i, J_i)_{1\le i\le 2}$ as coordinates of $\mathbb{S}^2_{r_1} \times \mathbb{S}^2_{r_2}$ near $\varepsilon(0,0,r_1,0,0,\pm r_2)$. The equations of motion are then given by

$$(\dot{M}_1, \dot{J}_1) = ((J_2 - d_{1,\gamma}J_1)L_1, (d_{1,\gamma}M_1 + M_2)L_1)$$

and

$$(\dot{M}_2, \dot{J}_2) = ((J_1 - d_{2,\gamma}J_2)L_2, (d_{2,\gamma}M_2 + M_1)L_2)$$

where $L_i = \pm r_i\sqrt{1-(M_i^2+J_i^2)/r_i^2}$. If linearized at $(0,0,\xi,0,0,\eta)$ where $\xi \in \{\pm r_1\}$, $\eta \in \{\pm r_2\}$, the corresponding linear system is given by the 4×4-matrix $A \equiv A_{\xi,\eta}$,

$$A = \begin{pmatrix} 0 & -d_{1,\gamma}\xi & 0 & \xi \\ d_{1,\gamma}\xi & 0 & \xi & 0 \\ 0 & \eta & 0 & -d_{2,\gamma}\eta \\ \eta & 0 & d_{2,\gamma}\eta & 0 \end{pmatrix}.$$

Let us compute $\det(\lambda - A) = \det(A - \lambda)$:

$$\det(A-\lambda) \,=\, \lambda^4 + a\lambda^2 + b^2, \tag{F.1}$$

where

$$a = d_{1,\gamma}^2\xi^2 + d_{2,\gamma}^2\eta^2 - 2\xi\eta \tag{F.2}$$

and

$$b = (d_{1,\gamma}d_{2,\gamma} - 1)\xi\eta. \tag{F.3}$$

The discriminant of (F.1) is given by

$$\begin{aligned} a^2 - 4b^2 &= (a-2b)(a+2b) \\ &= (d_{1,\gamma}^2\xi^2 - 2d_{1,\gamma}d_{2,\gamma}\xi\eta + d_{2,\gamma}^2\eta^2)(d_{1,\gamma}^2\xi^2 + 2d_{1,\gamma}d_{2,\gamma}\xi\eta + d_{2,\gamma}^2\eta^2 - 4\xi\eta). \end{aligned}$$

We first consider the case $\xi\eta = -r_1r_2$. Then

$$a^2 - 4b^2 = \underbrace{(d_{1,\gamma}r_1 + d_{2,\gamma}r_2)^2}_{\geq 0}\underbrace{((d_{1,\gamma}r_1 - d_{2,\gamma}r_2)^2 + 4r_1r_2)}_{>0} \geq 0$$

and

$$a = d_{1,\gamma}^2 r_1^2 + d_{2,\gamma}^2 r_2^2 + 2r_1r_2 > 0.$$

Hence $\lambda^2 = \mu_\pm$ where

$$\mu_\pm = -\frac{a}{2}\left(1 \pm \sqrt{1 - \left(\frac{2b}{a}\right)^2}\right) \leq 0. \tag{F.4}$$

More precisely, one always has $\mu_+ < 0$ whereas $\mu_- = 0$ iff $0 = a - 2b = (d_{1,\gamma}r_1 + d_{2,\gamma}r_2)^2$. Hence

$$\lambda_{1,2} = \pm i\sqrt{\frac{a}{2}}\left(1 + \sqrt{1 - \left(\frac{2b}{a}\right)^2}\right)^{\frac{1}{2}}$$

and

$$\lambda_{3,4} = \pm i\sqrt{\frac{a}{2}}\left(1 - \sqrt{1 - \left(\frac{2b}{a}\right)^2}\right)^{\frac{1}{2}}.$$

It follows that the fixed points $\pm(0,0,r_1,0,0,-r_2)$ of X_γ are both elliptic, except in the case $a - 2b = 0$ (i.e. $d_{1,\gamma}r_1 + d_{2,\gamma}r_2 = 0$) where they are degenerate elliptic.

Let us now turn to the case $\xi\eta = r_1r_2$. Then

$$\begin{aligned} a^2 - 4b^2 &= (d_{1,\gamma}r_1 - d_{2,\gamma}r_2)^2((d_{1,\gamma}r_1 + d_{2,\gamma}r_2)^2 - 4r_1r_2) \\ &= \left(\frac{r_1r_2}{s_{2k}}\right)^2\left(d_{1,\gamma}\sqrt{r} - d_{2,\gamma}\frac{1}{\sqrt{r}}\right)^2(f(\sqrt{r})^2 - 4s_{2k}^2), \end{aligned} \tag{F.5}$$

where we recall that $f(q) = (\gamma + s_k^2)q + (\gamma + c_k^2)/q$. First we have to establish some auxiliary results.

Lemma F.1. *Let $\gamma \in \mathbb{R}$ be arbitrary and assume that $1 \leq k < \frac{N}{4}$ and $0 < r < 1$. Then $d_{1,\gamma}\sqrt{r} - d_{2,\gamma}\frac{1}{\sqrt{r}} \neq 0$.*

Proof. Assume that $d_{1,\gamma}\sqrt{r} - d_{2,\gamma}\frac{1}{\sqrt{r}} = 0$. First note that if $d_{1,\gamma} = 0$, then $d_{2,\gamma} = 0$. Thus $d_{1,\gamma} = d_{2,\gamma}$ or $\gamma + s_k^2 = \gamma + c_k^2$. As $1 \leq k < \frac{N}{4}$ by assumption this leads to a contradiction. Hence $d_{1,\gamma} \neq 0$ and $d_{1,\gamma}\sqrt{r} - d_{2,\gamma}\frac{1}{\sqrt{r}} = 0$ is equivalent to $r = \frac{d_{2,\gamma}}{d_{1,\gamma}}$. As $r \leq 1$ by assumption we conclude that $d_{2,\gamma} \leq d_{1,\gamma}$ or $c_k^2 \leq s_k^2$ which contradicts the assumption $1 \leq k < \frac{N}{4}$. $\square$

In addition we will need the following

Lemma F.2. *Let $\gamma \in \mathbb{R}$ be arbitrary and assume that $1 \le k < \frac{N}{4}$ and $r > 0$. If $a < 0$, then $|f(\sqrt{r})| < 2s_{2k}$, where a is given by (F.2), i.e.*

$$\begin{aligned} a &= d_{1,\gamma}^2 r_1^2 + d_{2,\gamma}^2 r_2^2 - 2r_1 r_2 \\ &= \frac{r_1 r_2}{s_{2k}^2}((\gamma + s_k^2)^2 r + (\gamma + c_k^2)^2 \frac{1}{r} - 2s_{2k}^2) \end{aligned} \tag{F.6}$$

Proof. We argue indirectly and assume that for some γ_0, k, and r, we have $a < 0$ but $|f(\sqrt{r})| \ge s_{2k}$. The latter inequality can be written as

$$g_{\gamma_0}(r) := (\gamma_0 + s_k^2)^2 r + (\gamma_0 + c_k^2)^2 \frac{1}{r} \;\ge\; \frac{7}{2} s_{2k}^2 - 2\gamma_0^2 - 2\gamma_0. \tag{F.7}$$

In view of (F.6), the inequality $a < 0$ can be expressed as $g_{\gamma_0}(r) < 2s_{2k}^2$. Hence we may assume that the following inequalities hold

$$\frac{7}{2} s_{2k}^2 - 2\gamma_0^2 - 2\gamma_0 \le g_{\gamma_0}(r) < 2s_{2k}^2. \tag{F.8}$$

If $\gamma_0 = -s_k^2$ or $\gamma_0 = -c_k^2$, then one concludes

$$\frac{7}{2} s_{2k}^2 - 2\gamma_0^2 - 2\gamma_0 = 4s_{2k}^2,$$

contradicting (F.8). If $\gamma_0 \notin \{-s_k^2, -c_k^2\}$, then $g_{\gamma_0}(q) \to \infty$ for $q \to 0$ or $q \to \infty$. Further, for any $\gamma \in \mathbb{R} \setminus \{-s_k^2, -c_k^2\}$, g_γ achieves its minimum at $q = (\gamma + c_k^2)/(\gamma + s_k^2)$ and

$$\min_{q>0} g_\gamma(q) = 2\gamma^2 + 2\gamma + \frac{1}{2} s_{2k}^2. \tag{F.9}$$

If $\min_{q>0} g_\gamma(q) \ge \frac{7}{2} s_{2k}^2 - 2\gamma^2 - 2\gamma$, then by (F.9),

$$2\gamma^2 + 2\gamma \ge \frac{3}{2} s_{2k}^2. \tag{F.10}$$

Note that $\gamma \mapsto \min_{q>0} g_\gamma(q) = 2\gamma^2 + 2\gamma + \frac{1}{2} s_{2k}^2$ is a continuous function on $\mathbb{R}$, tending to ∞ as $\gamma \to \infty$. Hence there exists $\gamma_1 \in \mathbb{R}$ such that

$$g_{\gamma_0}(r) = \min_{q>0} g_{\gamma_1}(q).$$

In particular, $\min_{q>0} g_{\gamma_1}(q) > \min_{q>0} g_{\gamma_0}(q)$, i.e. $2\gamma_1^2 + 2\gamma_1 > 2\gamma_0^2 + 2\gamma_0$. When combined with (F.8) the latter inequality leads to

$$\begin{aligned} \frac{7}{2} s_{2k}^2 - 2\gamma_1^2 - 2\gamma_1 \;&\le\; \frac{7}{2} s_{2k}^2 - 2\gamma_0^2 - 2\gamma_0 \\ &= g_{\gamma_0}(r) = \min_{q>0} g_{\gamma_1}(q). \end{aligned}$$

By (F.10) we then have

$$2\gamma_1^2 + 2\gamma_1 \ge \frac{3}{2} s_{2k}^2. \tag{F.11}$$

On the other hand, as $\min_{q>0} g_{\gamma_1}(q) = g_{\gamma_0}(r) < 2s_{2k}^2$, one concludes from (F.9) that

$$2\gamma_1^2 + 2\gamma_1 + \frac{1}{2}s_{2k}^2 < 2s_{2k}^2. \tag{F.12}$$

Then (F.11) and (F.12) lead to the desired contradiction, and Lemma F.2 is proved. □

Proposition F.3. *Let $\gamma \in \mathbb{R}$ be arbitrary and assume that $1 \leq k < \frac{N}{4}$ and $0 < r \leq 1$. Then the following statements hold:*

(i) If $|f(\sqrt{r})| \geq 2s_{2k}$, then $\pm(0,0,r_1,0,0,r_2)$ are (possibly degenerate) elliptic fixed points of X_γ.

(ii) If $|f(\sqrt{r})| < 2s_{2k}$, then $\pm(0,0,r_1,0,0,r_2)$ are hyperbolic fixed points of X_γ. Their stable and and unstable manifolds have each dimension two.

Proof. In view of Lemma F.2 and formula (F.1) one concludes that the zeroes of $a^2 - 4b^2$ and $f(\sqrt{r})^2 - 4s_{2k}^2$ as well as the signs of the two expressions coincide.

(i) Assume that $f(\sqrt{r})^2 - 4s_{2k}^2 \geq 0$. We then conclude that $a^2 - 4b^2 \geq 0$. Further, by Lemma F.2, $a \geq 0$. In view of (F.1), the eigenvalues of A are then given by

$$\lambda_{1,2} = \pm(\mu_+)^{\frac{1}{2}}, \quad \lambda_{3,4} = \pm(\mu_-)^{\frac{1}{2}}$$

where

$$\mu_\pm = -\frac{a}{2} \pm \frac{1}{2}\sqrt{a^2 - 4b^2} \in \mathbb{R}.$$

As $a \geq 0$ it then follows that $\mu_\pm \leq 0$. Hence $(\lambda_i)_{1\leq i\leq 4}$ are purely imaginary, i.e. $\pm(0,0,r_1,0,0,r_2)$ are (possibly degenerate) elliptic fixed points of X_γ.

(ii) Assume that $f(\sqrt{r})^2 - 4s_{2k}^2 < 0$. Then

$$\mu_\pm = -\frac{a}{2} \pm \frac{i}{2}(d_{1,\gamma}r_1 - d_{2,\gamma}r_2)\sqrt{(d_{1,\gamma}r_1 + d_{2,\gamma}r_2)^2 - 4r_1r_2}.$$

By Lemma F.1 it then follows that Im $\mu_\pm \neq 0$, and we conclude that

$$\lambda_{1,2} = \pm(\mu_+)^{\frac{1}{2}}, \quad \lambda_{3,4} = \pm(\overline{\mu}_+)^{\frac{1}{2}}.$$

In particular, two eigenvalues have a positive real part and the other two a negative real part. Hence $\pm(0,0,r_1,0,0,r_2)$ are both hyperbolic fixed points of X_γ and the corresponding stable and unstable manifolds have each dimension two. □

Bibliography

[1] V. I. ARNOL'D, V. V. KOZLOV, & A. I. NEISHTADT, Mathematical Aspects of Classical and Celestial Mechanics. In: V. I. ARNOL'D (ED.), *Dynamical Systems II*, 2nd ed., Encyclopedia of Mathematical Sciences vol. 3, Springer, Berlin, 1993, 1-291.

[2] V. I. ARNOL'D, *Mathematical methods of Classical Mechanics*. 2nd ed., Graduate Texts in Mathematics **60**, Springer, New York, 1989.

[3] V. I. ARNOL'D, Proof of a theorem by A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Math. Surveys* **18** (1963), 9-36.

[4] V. I. ARNOL'D & A. AVEZ, *Ergodic Problems of Classical Mechanics*. W. A. Benjamin, New York, 1968.

[5] D. BAMBUSI & A. PONNO, Korteweg-de Vries equation and energy sharing in Fermi-Pasta-Ulam. *CHAOS* **15** (2005), 015107 (*5 pages*).

[6] D. BAMBUSI & A. PONNO, On Metastability in FPU. *Comm. Math. Phys.* **264** (2006), 539-561.

[7] G. P. BERMAN & F. M. IZRAILEV, The Fermi-Pasta-Ulam problem: Fifty years of progress. *CHAOS* **15** (2005), 015104 (*18 pages*).

[8] G. D. BIRKHOFF, *Dynamical Systems*. AMS Colloquium Publications, Vol. IX, 1927 [revised edition: 1966].

[9] H. W. BROER, KAM theory: the legacy of Kolmogorov's 1954 paper. *Bull. AMS (New Series)* **41**(4) (2004), 507-521.

[10] H. W. BROER, G. B. HUITEMA, & M. B. SEVRYUK, *Quasi-periodic Motions in Families of Dynamical Systems: Order Amidst Chaos*. Lecture Notes in Mathematics, Vol. 1645, Spinger, Berlin, 1996.

[11] T. BOUNTIS, H. SEGUR, & F. VIVALDI, Integrable Hamiltonian systems and the Painléve property. *Phys. Rev. A* **25**(3) (1982), 1257-1264.

[12] D. K. Campbell, P. Rosenau, & G. M. Zaslavsky, Introduction: The Fermi-Pasta-Ulam problem - The first fifty years. *CHAOS* **15** (2005), 015101 (*4 pages*).

[13] A. Carati, L. Galgani, & A. Giorgilli, The Fermi-Pasta-Ulam problem as a challenge for the foundations of physics. *CHAOS* **15** (2005), 015105 (*8 pages*).

[14] J. H. Conway & A. J. Jones, Trigonometric diophantine equations (On vanishing sums of roots of unity). *Acta Arithmetica* **XXX** (1976), 229-240.

[15] R. H. Cushman & L. M. Bates, *Global aspects of classical integrable systems*. Birkhäuser, Basel, 1997.

[16] T. Dauxois, M. Peyrard & S. Ruffo, The Fermi-Pasta-Ulam "numerical experiment": history and pedagogical perspectives. *Eur. J. Phys.* **26** (2005), 3-11.

[17] F. Fassò, M. Guzzo & G. Benettin, Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems. *Comm. Math. Phys.* **197** (1998), 347-360.

[18] H. Flaschka, The Toda lattice. I. Existence of integrals. *Phys. Rev.*, Sect. B **9** (1974), 1924-1925.

[19] E. Fermi, Beweis, dass ein mechanisches Normalsystem im allgemeinen quasi-ergodisch ist. *Phys. Zeitschrift* **24** (1923), 261-265.

[20] E. Fermi, J. Pasta & S. Ulam, Studies of non linear problems. *Los Alamos Rpt.* **LA-1940** (1955). In: *Collected Papers of Enrico Fermi*. University of Chicago Press, Chicago, 1965, Volume II, 978-988. Theory, Methods and Applications, 2nd ed., Marcel Dekker, New York, 2000. Also contained in: S. M. Ulam, *Analogies between Analogies. The mathematical reports of S. M. Ulam and his Los Alamos collaborators*. Edited by A. R. Bednarek & F. Ulam. University of California Press, Berkeley, 1990.

[21] J. Ford, Equipartition of energy for nonlinear systems. *J. Math. Phys.* **2**(3) (1961), 387-393.

[22] J. Ford & J. Waters, Computer studies of energy sharing and ergodicity for nonlinear oscillator systems. *J. Math. Phys.* **4**(10) (1963), 1293-1306.

[23] J. Ford, The Fermi-Pasta-Ulma problem: paradox turns discovery. *Physics Reports* **213**(5) (1992), 271-310.

[24] F. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo & A. Vulpiani, Approach to equilibrium in a chain of nonlinear oscillators. *J. de Physique* **43** (1982), 707-713.

[25] P. Galison, *How Experiments End*. University of Chicago Press, Chicago, 1987.

[26] C. S. Gardner, J. M. Greene, M. D. Kruskal & R. M. Miura, Method for solving the Korteweg-de Vries equation. *Phys. Rev. Lett.* **19**, 1095-1097 (1967).

[27] C. S. Gardner, J. M. Greene, M. D. Kruskal & R. M. Miura, Korteweg-de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. *J. Math. Phys.* **9**, 1204-1209 (1968).

[28] B. Grammaticos, A. Ramani & V. Papagiorgiou, Do Integrable Mappings Have the Painléve Property? *Phys. Rev. Lett.* **67**(14) (1991), 1825-1828.

[29] F. G. Gustavson, On constructing formal integrals of a Hamiltonian system near an equilibrium point. *Astronom. J.* **71**(8), 670-686 (1966).

[30] M. Hénon, Integrals of the Toda lattice. *Phys. Rev.*, Sect. B **9** (1974), 1921.

[31] M. Hénon & C. Heiles, The applicability of the third integral of motion: Some numerical experiments. *Astronom. J.* **69**(1), 73-79 (1964).

[32] A. Henrici & T. Kappeler, Global action-angle variables for the periodic Toda lattice. In preparation.

[33] A. Henrici & T. Kappeler, Global Birkhoff coordinates for the periodic Toda lattice. In preparation.

[34] A. Henrici & T. Kappeler, Birkhoff normal form for the periodic Toda lattice. `arXiv:nlin/0609045v1 [nlin.SI]`. To appear in *Contemp. Math.*

[35] A. Henrici & T. Kappeler, Results on normal forms for FPU chains. `arXiv:nlin/0611063v2 [nlin.SI]`. *Comm. Math. Phys.* **278** (2008), 145-177.

[36] A. Henrici & T. Kappeler, Resonant normal form for even periodic FPU chains. `arXiv:0709.2624v1 [nlin.SI]`.

[37] Yu. S. Il'yaschenko, A steepness test for analytic functions. *Uspekhi Mat. Nauk* **41** (1986), 193-194 [Russian]. English translation in: *Russian math. Surveys* **41** (1986), 229-230.

[38] F. M. Izrailev & B. V. Chirikov, Statistical properties of a nonlinear string. *Doklady Akademii Nauk SSSR* **166**(1) (1965) [Russian]. English translation in: *Soviet Phys. Dokl.* **11**(1) (1966), 30-32.

[39] E. A. Jackson, Nonlinear coupled oscillators I: Perturbation theory; ergodic problem. *J. Math. Phys.* **4**(4) (1963), 551-558.

[40] E. A. Jackson, Nonlinear coupled oscillators II: Comparison of theory with computer simulations. *J. Math. Phys.* **4**(5) (1963), 686-700.

[41] R. Jost, Winkel- und Wirkungsvariable für allgemeine mechanische Systeme. *Helvetica Physica Acta* **41** (1968), 965-968.

[42] T. Kappeler & J. Pöschel, *KdV & KAM*. Ergebnisse der Mathematik, 3. Folge, **45**, Springer, Berlin, 2003.

[43] A. N. Kolmogorov, On the conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR* **98** (1954), 527-530 [Russian]. English translation in: *Lecture Notes in Physics* **45**, Springer, 1979.

[44] D. J. Korteweg, G. de Vries, On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Phil. Mag. Ser. 5* **39**, 422-443 (1895).

[45] T. Kuhn, *The Structure of Scientific Revolutions*. 3rd ed., University of Chicago Press, Chicago, 1996.

[46] S. Lang, *Fundamentals of Differential Geometry*. Graduate Texts in Mathematics **191**, Springer, New York, 1999.

[47] P. D. Lax, Integrals of nonlinear equations of evolution and solitary waves. *Comm. Pure Appl. Math.* **28**, 141-188 (1975).

[48] R. de la Llave, A tutorial on KAM theory. *Proc. Symp. Pure Math.* **69**, 175-292 (2001).

[49] P. Lochak, Hamiltonian perturbation theory: periodic orbits, resonances and intermittency. *Nonlinearity* **6** (1993), 885-904.

[50] T. Lucretius Carus, *On the Nature of the Universe* (Verse translation by James H. Mantinband). Ungar, New York, 1965.

[51] S. V. Manakov, Complete integrability and stochastization of discrete dynamical systems. *Zh. Exp. Teor. Fiz.* **67** (1974), 543-555 [Russian]. English translation: *Sov. Phys. JETP* **40** (1975), 269-274.

[52] L. Markus & K. Meyer, Generic Hamiltonian systems are neither integrable nor ergodic. *Mem. Am. Math. Soc.* **45**, 1974.

[53] J. Moser, On invariant curves of area preserving mappings of an annulus. *Nachr. Akad. Wiss. Gött., Math. Phys. Kl.* (1962), 1-20.

[54] N. N. Nekhoroshev, Behaviour of Hamiltonian systems close to integrable. *Funct. Anal. Appl.* **5** (1971), 338-339.

[55] N. N. Nekhoroshev, An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems I. *Uspekhi Mat. Nauk* **32** (1977), 5-66 [Russian]. English translation: *Russian Math. Surveys* **32** (1977), 1-65.

[56] N. N. Nekhoroshev, An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems II. *Trudy Sem. Petrovsk* **5** (1979), 5-50 [Russian]. English translation: *Topics in Modern Mathematics, Petrovskii Sem. No. 5* O. A. Oleinik Ed., Consultant Bureau, New York, 1985.

[57] L. Niederman, Nonlinear stability around an equilibrium point in a Hamiltonian system. *Nonlinearity* **11** (1998), 1465-1479.

[58] T. Nishida, A note on an existence of conditionally periodic oscillation in a one-dimensional lattice. *Mem. Fac. Engrg. Kyoto Univ.* **33** (1971), 27-34.

[59] H. Poincaré, *Les Méthodes Nouvelles de la Méchanique Céleste*, Vol. I-III, Gauthier-Villars, Paris, 1899.

[60] M. Petkovšek, H. S. Wilf, & D. Zeilberger, *A=B*, A. K. Peters, Wellesley, 1996.

[61] M. Pettini, L. Casetti, M. Cerruti-Sola, R. Franzosi, & E. G. D. Cohen, Weak and strong chaos in Fermi-Pasta-Ulam models and beyond. *CHAOS* **15** (2005), 015106 (*13 pages*).

[62] J. Pöschel, Integrability of Hamiltonian Systems on Cantor Sets. *Comm. Pure Appl. Math.* **35** (1982), 653-695.

[63] J. Pöschel, Nekhoroshev Estimates for Quasi-Convex Hamiltonian Systems. *Math. Z.* **213** (1993), 187-216.

[64] J. Pöschel, On Nekhoroshev's Estimate at an Elliptic Equilibrium. *Int. Math. Res. Not.* **4** (1999), 203-215.

[65] J. Pöschel, A Lecture on the Classical KAM Theorem. *Proc. Symp. Pure Math.* **69** (2001), 707-732.

[66] P. Poggi & S. Ruffo, Exact solutions in the FPU oscillator chain. *Physica D* **103** (1997), 251-272.

[67] A. P. Prudnikov, Yu. A. Brychkov & O. I. Marichev, *Integrals and Series, Vol. 1: Elementary Functions*, Gordon and Breach, New York, 1986.

[68] A. F. Ranada, A. Ramani, B. Dorizzi, & B. Grammaticos, The weak-Painléve property as a criterion for the integrability of dynamical systems. *J. Math. Phys.* **26**(4) (1985), 708-710.

[69] B. Rink & F. Verhulst, Near-integrability of periodic FPU-chains. *Physica A* **285** (2000), 467-482.

[70] B. Rink, Symmetry and resonance in periodic FPU chains. *Comm. Math. Phys.* **218** (2001), 665-685.

[71] B. Rink, Direction reversing travelling waves in the Fermi-Pasta-Ulam chain. *J. Nonlinear Sci.* **12** (2002), 479-504.

[72] B. Rink, Geometry and dynamics in Hamiltonian lattices. *Thesis,* Universiteit Utrecht, 2003.

[73] B. Rink, Proof of Nishida's conjecture on anharmonic lattices. *Comm. Math. Phys.* **261** (2006), 613-627.

[74] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Reg. Chaot. Dyn.* **6** (2001), 119-204.

[75] J. A. Sanders, *On the theory of nonlinear resonance*, Thesis, University of Utrecht, Utrecht, 1979.

[76] G. Séroul, *Programming for Mathematicians*, Springer, Berlin, 2000.

[77] M. B. Sevryuk, The finite-dimensional reversible KAM theory. *Physica D* **112** (1998) (special issue on "Time-Reversal Symmetry in Dynamical Systems"), 132-147.

[78] G. Teschl, Almost Everything You Always Wanted to Know About the Toda Equation. *Jber. d. Dt. Math.-Verein.* **103** (2001), 149-162.

[79] M. Toda, *Theory of Nonlinear Lattices*, 2nd enl. ed., Springer Series in Solid-State Sciences **20**, Springer, Berlin, 1989.

[80] S. Ulam, *A collection of mathematical problems*, Interscience Inc., New York, 1960.

[81] B. L. Van der Waerden, *Algebra I*, Heidelberger Taschenbücher, Springer, Berlin, 1966.

[82] B. C. van Frassen, *The Scientific Image*, Clarendon Press, Oxford, 1980.

[83] G. H. Walker & J. Ford, Amplitude instability and ergodic behavior for conservative nonlinear oscillator systems. *Phys. Rev.* **188**(1), 416-432 (1969).

[84] T. P. Weissert, *The genesis of simulation in dynamics: pursuing the Fermi-Pasta-Ulam problem*, Springer, New York, 1997.

[85] N. J. Zabusky, Exact solution for the vibrations of a nonlinear continuous model string. *J. Math. Phys.* **3**(5), 1028-1039 (1962).

[86] N. J. Zabusky, Phenomena associated with the oscillations of a nonlinear model string: The problem of Fermi, Pasta, Ulam. In: S. Drobot & P. A. Viebrock (Eds.), *Mathematical Models in Physical Science*, Proceedings of a Conference at the University of Notre Dame, 1962, Prentice-Hall, Englewood Cliffs, 1962, 99-133.

[87] N. J. Zabusky & M. D. Kruskal, Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**(6), 240-243 (1965).

[88] N. J. Zabusky, Fermi-Pasta-Ulam, solitons and the fabric of nonlinear and computational science: History, synergetics, and visiometrics. *CHAOS* **15** (2005), 015102 (*16 pages*).

[89] G. M. Zaslavsky, Long way from the FPU-problem to chaos. *CHAOS* **15** (2005), 015103 (*10 pages*).

[90] E. Zeidler (Hrsg.), *Teubner-Taschenbuch der Mathematik*, 2., durchgesehene Auflage. Teubner, Wiesbaden, 2003.

[91] N. T. Zung, Kolmogorov condition near hyperbolic singularities of integrable Hamiltonian systems. `arXiv:0706.1590v2 [math.DS]`.

Acknowledgements

First of all, I would like to thank my advisor, Thomas Kappeler, for his advice during all the years we have been working together. It is really exceptional how much time and energy he has devoted to guiding me through my work, how many (erroneous) drafts of my papers he has meticulously examined, how much patience he brought up in guiding me through the subtleties of our research area and introducing me to new viewpoints. Especially the more geometric ones are a great amplification of my a priori more analytic viewpoint - the way the moment map analysis is done can be seen as an example of this.

Many thanks also to Jürgen Pöschel for refereeing my thesis - needless to say that I have also enormously profited from his joint work with Thomas Kappeler on the KdV equation.

It was a great pleasure to discuss my work with Bob Rink. I have greatly profited from his previous work on the same subject, which gave my a lot of hints and insight into what can be accomplished in this area, and our various discussions in Dresden, Zürich, and Leiden (and the electronic ones) have given me deeper insight into his work, from which my work has again profited.

Many thanks also to our collaborators in Milano, in particular Luigi Galgani, Antonio Giorgilli, Dario Bambusi, Simone Paleari, and Tiziano Penati, with whom I am greatful for having had the opportunity to present my work at their seminar in Milano, and I am very curious what will come out of our ongoing collabarations concerning the numerical implementation of the results of this thesis.

I am deeply indepted to Percy Deift, who showed great interest in my work and took the time to considerably improve some of my nondegeneracy computations - that they were based on wrong assumptions is of course my fault.

It was a wonderful experience for me to spend a semester at Humboldt-Universität zu Berlin - many thanks to Jürg Kramer and all other members of the Graduiertenkolleg “Arithmetik und Geometrie” for their hospitality. I would also like to thank Jochen Brüning for showing a lot of interest in my work and giving me the opportunity to present some of my results in his research seminar. Thanks also to Nicolas Roy and Evgeny Korotyaev for discussing with me my research.

Finally, thanks to Marie-Louise Henrici for proofreading the thesis and to Gerda Schacher, Carsten Rose, Bruno Nietlispach, and Mirko Birbaumer for their help with the layout.

Curriculum Vitae

- **Name**: HENRICI
- **Vornamen**: Andreas Dominik
- **Geburtsdatum**: 25. Juli 1977
- **Heimatort**: Basel BS
- **Ausbildung**:
  - 1989-1996: Kantonsschule Hohe Promenade Zürich, *Abschluss*: Matura Typus A (1996).
  - 1993-1999: Konservatorium Zürich, Violoncello-Studium, *Abschluss*: Lehr- und Orchesterdiplom (1999).
  - 1998-2003: ETH Zürich, Studium der Mathematik, *Abschluss*: Diplom mit Auszeichnung (2003).
  - Diplomfach: Mathematik
  - Diplomarbeit: "Normal Form Theory for the Periodic Toda Lattice", betreut von Prof. Dr. Thomas Kappeler (Universität Zürich) und Prof. Dr. Eduard Zehnder (ETH Zürich)
  - 2004-2007: Universität Zürich/Humboldt-Universität zu Berlin, Doktorat in Mathematik
- **Gegenwärtige Stellung**: Assistent am Institut für Mathematik der Universität Zürich (seit Oktober 2003)

Die vorliegende Dissertation, "Normal Form for Fermi-Pasta-Ulam Chains", entstand unter der Leitung von Prof. Dr. Thomas Kappeler.